

On the Continuity of Characteristic Functionals and Sparse Stochastic Modeling

Julien Fageot, Arash Amini, and Michael Unser

January 29, 2014

Abstract. The characteristic functional is the infinite-dimensional generalization of the Fourier transform for measures on function spaces. It characterizes the statistical law of the associated stochastic process in the same way as a characteristic function specifies the probability distribution of its corresponding random variable. Our goal in this work is to lay the foundations of the *innovation model*, a (possibly) non-Gaussian probabilistic model for sparse signals. This is achieved by using the characteristic functional to specify sparse stochastic processes that are defined as linear transformations of general continuous-domain white Lévy noises (also called *innovation processes*). We prove the existence of a broad class of sparse processes by using the Minlos-Bochner theorem. This requires a careful study of the regularity properties, especially the L^p -boundedness, of the characteristic functional of the innovations. We are especially interested in the functionals that are only defined for $p < 1$ since they appear to be associated with the sparser kind of processes. Finally, we apply our main theorem of existence to two specific subclasses of processes with specific invariance properties.

Keywords. Characteristic functional - Generalized stochastic process - Innovation model - White Lévy noise.

Mathematics Subject Classification. 60G20 - 60H40 - 60G18.

1 Introduction

1.1 Presentation of the Innovation Model

Sparsity plays a significant role in the mathematical modeling of real-world signals. A signal is said to be *sparse* if its energy tends to be concentrated in few coefficients in some transform-domain. Natural images are known to have such a sparse representation. Moreover, numerous statistical studies have shown that typical biomedical and natural images are non-Gaussian [SLSZ03]. These empirical facts highlight the fundamental limits of probabilistic models based on Gaussian priors [MG01]. The sparsity-based theories developed for overcoming these limitations include wavelets [Mal99] (with powerful applications in image coding and processing) and, more recently, compressed sensing [Don06, CRT06]. They are inherently deterministic.

A new general model has been recently developed in order to reconcile the sparsity paradigm of signal processing with a probabilistic formulation. Its general foundations and motivations were discussed in [UTS11, UTAK11]. The main hypotheses are as follows.

- A signal is modeled as a random continuous-domain function s defined on \mathbb{R}^d . Hence, s is the stochastic process that captures the statistical properties of the signal.
- The process s can be linearly decoupled, which implies the existence of a linear whitening operator L such that

$$Ls = w, \tag{1}$$

where w is a continuous-domain innovation process, also called white noise, which is *not necessarily Gaussian*. The term “innovation” reflects the property that w is the unpredictable component of the process.

Following the terminology of [UTS11], these two hypotheses define the *innovation model* (see Figure 1).

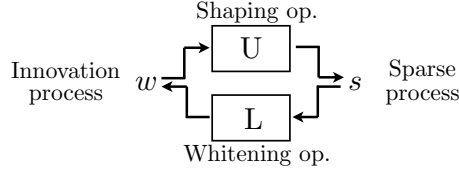


Figure 1: Innovation model

The innovation model provides a mathematical framework that complies with the sparse behavior of real-world signals for at least two theoretical reasons. First, a process s following (1), given a non-Gaussian innovation process w , has been shown to be sparser (*i.e.*, more compressible) than any Gaussian one [UTS11]. We can therefore refer to these processes as *sparse processes*. Second, it has been demonstrated for the case of symmetric α -stable (S α S) AR(1) processes that better decoupling is achieved in a wavelet-like representation than with the traditional sine basis or the KLT [PU13]. The innovation model has already been applied to various fields of image processing such as Bayesian estimation from noisy samples of sparse processes [AKBU13], algorithms for the optimal quadratic estimation of sparse processes [KPAU13], and reconstruction techniques based on sparse and self-similar processes [BFKU13]. It was also found to be useful in inverse problems, involving MRI, deconvolution, and X-ray tomography reconstruction problems [BKNU13, BKU12]. While these examples show that sparse processes are highly relevant for practical applications, the theory currently available is based on too-constraining assumptions. In particular, it excludes some of the sparsest processes such as S α S with $\alpha < 1$, for which wavelets have been found empirically to be optimal [PU13]. More generally, the compatibility between a linear operator L and an innovation process w , defined as the existence of a process s such that $Ls = w$, is a crucial question that needs to be addressed.

The innovation model is formulated within the extended framework of Generalized Stochastic Processes (GSP), the stochastic counterpart of Schwartz theory of generalized functions. This probabilistic theory was historically introduced in the 50's by Gelfand [Gel55]. The larger part of the literature on GSP is concerned with second-order processes with a special emphasis on the Gaussian case. Recent examples include results on existence and regularity of Gaussian GSP [BORV98], construction of Gaussian and Poisson isotropic and self-similar GSP [BEK10], and classification of Gaussian stationary mean-square increments GSP [AJ12]. Gelfand's formulation extends beyond the second-order family. For instance, the processes with unbounded variance take a particular relevance in the context of sparsity [AUM11], [GCD12]. We consequently consider [GV64] as the starting point of our own developments.

This framework of generalized stochastic processes enables the definition of the innovation processes, which cannot be defined as classical stochastic processes. Gelfand defines innovation processes as random generalized functions. As a consequence, the GSP are not observed pointwise but by forming duality products with test functions $\varphi \in \mathcal{D}$, where \mathcal{D} is the space of smooth and compactly supported functions. For a fixed φ , the observation $\langle w, \varphi \rangle$ is then a conventional random variable. In addition to that, Gelfand's framework appears to be particularly adapted for the development of the theory related to the innovation model and for its applicability in signal processing.

1.2 Contributions

Our goal is to define the broadest framework that guarantees the existence of sparse processes. In that sense, our work can be seen as an extension of the existence results of [UTS11]. We now give the three main results of this paper.

a) Proper definition of innovation processes over \mathcal{S}'

While the usual definition of an innovation process (*i.e.*, a continuous-domain white noise) is over the space \mathcal{D}' , the topological dual of \mathcal{D} [GV64], we therein define innovation processes over the space \mathcal{S}'

of tempered generalized functions. This extension requires to identify a sufficient condition for an innovation process to be tempered, which is the focus of Theorem 3. We show in addition that most of the innovation processes in \mathcal{D}' are supported on \mathcal{S}' (Theorem 4). This ensures the compatibility between the two constructions over \mathcal{D}' and \mathcal{S}' when they occur.

The choice of \mathcal{S}' is driven by the desire to make the innovation model applicable to signal processing. Adopting \mathcal{S}' as the definition space allows us to extend $\langle w, \varphi \rangle$ to the case of non-compactly supported functions, which are crucial in signal-processing applications.

b) Existence of the broader family of sparse processes

We investigate the compatibility of pairs (w, L) of innovation processes and linear operators and introduce a large class of valid combinations. Before describing our contributions, we briefly summarize the known results on the existence of sparse processes of the form $s = L^{-1}w$.

Gelfand formulates the general definition of innovation processes on \mathcal{D}' . Hence, (w, Id) is a compatible pair for all w defined by [GV64]. An immediate extension of this result is that (w, L) is also valid if the adjoint operator L^{-1} has a \mathcal{D} -stable inverse operator. In that case, one can directly define s according to

$$\langle s, \varphi \rangle = \langle w, L^{*-1} \varphi \rangle. \quad (2)$$

Indeed, the \mathcal{D} -stability of L^{*-1} ensures that $\langle w, L^{*-1} \varphi \rangle$ is always well-defined. We then have $\langle Ls, \varphi \rangle = \langle s, L^* \varphi \rangle = \langle w, L^{*-1} L^* \varphi \rangle = \langle w, \varphi \rangle$ or, equivalently, $Ls = w$. Unfortunately, interesting whitening operators for signal processing do not fulfill this stability condition. For instance, in the one-dimension case, the common differential operator $L = \alpha \text{Id} - D$ with $\alpha > 0$ associated with AR(1) processes is already problematic. Indeed, $\rho_\alpha(t) = u(t)e^{-\alpha t}$, with $u(t)$ the Heaviside step function, is the causal Green function of $L^* = \alpha \text{Id} + D$. Its inverse $L^{*-1} \varphi = (\rho * \varphi)(t)$ is therefore not \mathcal{D} -stable. Note that L^{*-1} is however \mathcal{S} -stable. This has also encourage us to develop the theory of innovation model over \mathcal{S}' instead of \mathcal{D}' .

As a first step, Unser et al. have expressed the comparability condition over \mathcal{S}' . The pair (w, L) is shown to be compatible if there exists $p \geq 1$ such that (i) L^* admits a left-inverse operator L^{*-1} from \mathcal{S} to L^p and (ii) w is p -admissible, a condition that quantifies the level of sparsity of w (see Theorem 3 in [UTS11]). This theory enables sparse processes to be defined not only for classical differential operators that are typically \mathcal{S} -stable, but also for fractional differential operators which require the L^p extension. This generalization is sufficient for most of the cases of practical interest, but has two restrictions. First, it does not encompass the case of high sparsity as it is restricted to $p \geq 1$. Second, the p -admissibility condition limits the generality of the results.

In this paper, we formulate a new criterion of compatibility that avoid these restrictions. We show that the generalized stochastic process s over \mathcal{S}' with $Ls = w$ exists if one can link the existence of moments for w to the existence of a stable left-inverse for the operator L^* (Theorem 5). We present our proofs of sufficiency in the most general setting, which requires us to extend the continuity of the characteristic functional of the innovation processes from \mathcal{S} to L^p spaces (Proposition 4).

c) Construction of specific subclasses of processes

We apply our compatibility conditions to two specific families of operators. A class of self-similar processes is defined (Proposition 6) by extending a previous study of the fractional Laplacian operators [SU12]. A class of directional Lévy processes in dimension d is introduced by the use of directional differential operators (Proposition 9). The latter extends the work done in [UTS11] for $d = 1$.

1.3 Outline

This paper is organized as follows. In Section 2, we recall some concepts and results on generalized stochastic processes. In Section 3, we first present the general construction of white noises (innovation processes) developed by Gelfand [GV64] and adapt it to the space of tempered generalized functions (Section 3.1). Next, we present and prove a criterion for the compatibility of the innovation process w and the whitening operator L to form a sparse process (Section 3.2). The proof relies on continuity bounds for the characteristic functional of innovation processes (Section 3.3). Finally, we apply this criterion in Section 4 to two specific classes of operators and identify classes of generalized stochastic processes: self-similar sparse processes through fractional Laplacian operators (Section 4.1) and directional sparse processes through directional-derivative operators (Section 4.2).

2 Generalized Stochastic Processes

Our main concern is to define stochastic processes that satisfy the innovation model (1). The theory of generalized stochastic processes is based on functional analysis. In Table 1, we provide the definition of function spaces linked to our work. They include subspaces of ordinary functions from \mathbb{R}^d to \mathbb{R} (classical functions) as well as subspaces of the space \mathcal{D}' of distributions (also called generalized functions) [Sch66].

2.1 Definition of Generalized Stochastic Processes

We deviate from the traditional time-series approach to stochastic processes by presenting them as probability measures on a function space \mathcal{X} of functions from E to \mathbb{R} . Let \mathcal{X} be a topological vector space of real-valued functions. We denote by \mathcal{A} the σ -field generated by the cylindric sets. There are the subsets of E defined by $A_{\mathbf{x},B} = \{h \in \mathcal{X}, (h(x_1), \dots, h(x_N)) \in B\}$ for fixed $N \in \mathbb{N}$, where $\mathbf{x} = (x_1, \dots, x_N) \in E^N$ and B is a Borelian set in \mathbb{R}^N . For a given probability measure \mathcal{P} on \mathcal{A} , the *canonical stochastic process* s on $(\mathcal{X}, \mathcal{A}, \mathcal{P})$ is defined by

$$\begin{aligned} s : (\mathcal{X}, E) &\rightarrow \mathbb{R} \\ (h, x) &\mapsto h(x). \end{aligned}$$

There are two ways to consider s .

- If $\mathbf{x} = (x_1, \dots, x_N) \in E^N$ is fixed, $h \mapsto (h(x_1), \dots, h(x_N))$ is a random variable in \mathbb{R}^N with probability law $\mathcal{P}_{\mathbf{x}}(B) = \mathcal{P}(h \in \mathcal{X}, (h(x_1), \dots, h(x_N)) \in B)$ for any Borelian set B of \mathbb{R}^N . These laws are the *finite-dimensional marginals* of s .
- For $h \in \mathcal{X}$ following the probability measure \mathcal{P} , the mapping $x \mapsto h(x)$ is a sample function of the stochastic process (*i.e.*, a random element of \mathcal{X}).

If $E = \mathbb{R}^d$, we get back to the theory of classical (non-generalized) stochastic processes. The generalized theory of stochastic processes is obtained when E is formed by a set of test functions of \mathbb{R}^d to \mathbb{R} . Let $E = \mathcal{T}$ be a locally convex topological vector space (l.c.t.v.s.)—the minimal structure required in functional analysis [Rud91]—and let $\mathcal{X} = \mathcal{T}'$ be the topological dual of \mathcal{T} . We define the stochastic process s by

$$\begin{aligned} s : (\mathcal{T}', \mathcal{T}) &\rightarrow \mathbb{R} \\ (u, \varphi) &\mapsto \langle u, \varphi \rangle. \end{aligned} \tag{3}$$

The random variable $s(\cdot, \varphi)$ is denoted by $\langle s, \varphi \rangle$. The realization $s(u, \cdot)$, which follows \mathcal{P} for $u \in \mathcal{T}'$, is by definition a linear and continuous functional on \mathcal{T} . We call such s a *generalized stochastic process* if $\mathcal{D} \subset \mathcal{T} \subset L^2 \subset \mathcal{T}' \subset \mathcal{D}'$, meaning that s is a random generalized function. In [GV64], Gelfand and Vilenkin develop the essential results for $\mathcal{T} = \mathcal{D}$. In this paper, we especially focus on its extensions to $\mathcal{T} = \mathcal{S}$.

In Table 1, we list the type of structures that are useful in this paper. All considered function spaces are vectorial and most of them are locally convex, as they have a topological structure inherited from a distance, a norm, or a family of semi-norms [Rud91]. There are essentially two main structures for function spaces: Banach structure for complete normed spaces and nuclear structure for spaces defined by a family of semi-norms that imply suitable decreasing properties on unit balls [Tre67]. Note that the Banach and nuclear structures are mutually exclusive in infinite dimension. As we shall see, the nuclear structure is central to the definition of a continuous-domain innovation process.

2.2 The Characteristic Functional

By analogy with the finite-dimensional case ($\mathcal{T} = \mathbb{R}^N$), where the characteristic function characterizes a probability measure (Lévy's theorem), we use a Fourier-domain representation to describe the measures on \mathcal{T}' .

Table 1: Definition of function spaces used in the paper

Space	Parameter	Definition	Structure
L^p	$1 \leq p < +\infty$	$\ f\ _p = (\int_{\mathbb{R}^d} f(\mathbf{r}) ^p d\mathbf{r})^{1/p} < +\infty$	Complete normed
L^∞		$\ f\ _\infty = \text{ess sup}_{\mathbf{r} \in \mathbb{R}^d} f(\mathbf{r}) < +\infty$	Complete normed
L^p	$0 < p < 1$	$d_p(f, g) = \int_{\mathbb{R}^d} (f - g)(\mathbf{r}) ^p d\mathbf{r} < +\infty$	Complete metric
$L^{\infty, \alpha}$	$0 < \alpha < +\infty$	$\ f\ _{\infty, \alpha} = \ (1 + \ \mathbf{r}\ _2^\alpha) f(\mathbf{r})\ _\infty < +\infty$	Complete normed
\mathcal{R}		$\bigcap_{\alpha > 0} L^{\infty, \alpha}$	Complete metric
C^∞		Infinitely differentiable functions	Vectorial
\mathcal{D}		$f \in C^\infty$ with compact support	Nuclear
\mathcal{S}		$f \in C^\infty$ with $\partial^{\mathbf{n}} f \in \mathcal{R}$ for all $\mathbf{n} \in \mathbb{N}^d$	Nuclear
\mathcal{O}_M		Space of slowly increasing functions s.t. $f \in C^\infty$ with $ \partial^{\mathbf{n}} f(\mathbf{r}) \leq P_{\mathbf{n}}(\mathbf{r}) $ for some polynomial $P_{\mathbf{n}}$ and all \mathbf{n}	Vectorial
\mathcal{D}'		u linear and continuous functional on \mathcal{D}	Nuclear
\mathcal{S}'		u linear and continuous functional on \mathcal{S}	Nuclear
\mathcal{O}'_C		Space of rapidly decreasing generalized functions or equivalently $u \in \mathcal{S}'$ such that $\mathcal{F}u \in \mathcal{O}_M$	Vectorial

Definition 1. Let \mathcal{T} be a l.c.t.v.s. and let \mathcal{T}' be its topological dual. The characteristic functional of a generalized stochastic process s on \mathcal{T}' associated with the probability measure \mathcal{P}_s is defined as

$$\widehat{\mathcal{P}}_s(\varphi) = \mathbb{E} \left[e^{j\langle s, \varphi \rangle} \right] = \int_{\mathcal{T}'} e^{j\langle u, \varphi \rangle} d\mathcal{P}_s(u), \quad (4)$$

where $\varphi \in \mathcal{T}$.

The characteristic functional contains the definition of all finite-dimensional laws of the process, in particular the distribution of all random vectors $X = (\langle s, \varphi_1 \rangle, \dots, \langle s, \varphi_N \rangle)$. Indeed, the characteristic function of X is given by

$$\widehat{p}_X(\boldsymbol{\omega}) = \mathbb{E}[e^{j\langle \boldsymbol{\omega}, X \rangle}] = \widehat{\mathcal{P}}_s(\omega_1 \varphi_1 + \dots + \omega_N \varphi_N). \quad (5)$$

In Proposition 1, we summarize the main properties of $\widehat{\mathcal{P}}_s$.

Proposition 1. A characteristic functional is normalized ($\widehat{\mathcal{P}}_s(0) = 1$) and is continuous and positive-definite on \mathcal{T} . The latter means that for all $N \in \mathbb{N}$, $a_1, \dots, a_N \in \mathbb{C}$, and $\varphi_1, \dots, \varphi_N \in \mathcal{T}$, we have that

$$\sum_{i,j} a_i \bar{a}_j \widehat{\mathcal{P}}_s(\varphi_i - \varphi_j) \geq 0. \quad (6)$$

The normalization property reflects the fact that $\mathcal{P}_s(\mathcal{T}) = 1$, whereas the positive-definiteness is linked with the non-negativity of the measure \mathcal{P}_s . Our focus here is on probability measures on the dual space \mathcal{N}' of a nuclear space $\mathcal{N} \subset L^2 \subset \mathcal{N}'$. The reason is that the converse of Proposition 1 also holds if \mathcal{T} is nuclear (Theorem 1). Notorious examples of nuclear spaces are \mathcal{D} , \mathcal{S} , and their duals \mathcal{D}' (the space of distributions) and \mathcal{S}' (the space of tempered distributions), as seen in Table 1. This highlights the deep link between nuclear structures and the theory of generalized processes.

Theorem 1 (Minlos-Bochner). Let \mathcal{N} be a nuclear space and $\widehat{\mathcal{P}}$ be a continuous, positive-definite functional from \mathcal{N} to \mathbb{C} with $\widehat{\mathcal{P}}(0) = 1$. Then, there exists a unique measure \mathcal{P}_s on \mathcal{N}' such that

$$\widehat{\mathcal{P}} = \widehat{\mathcal{P}}_s.$$

Minlos-Bochner's theorem is an extension of Bochner's theorem to the infinite-dimensional setting. It is our key tool to define probability measures on the nuclear space \mathcal{S}' .

3 A Criterion for Existence of Sparse Processes

3.1 White Lévy-Schwartz Noise

We first recall some definitions and results from Gelfand and Vilenkin's theory of generalized stochastic processes [GV64], especially the definition of white Lévy noises on \mathcal{D} .

Innovation Processes on \mathcal{D}'

Definition 2. A stochastic process w on \mathcal{D}' characterized by the probability measure \mathcal{P}_w is said to be

- with independent value at every point if the random variables $X_1 = \langle w, \varphi_1 \rangle$ and $X_2 = \langle w, \varphi_2 \rangle$ are independent whenever $\varphi_1, \varphi_2 \in \mathcal{D}$ have disjoint supports (i.e., if $\varphi_1 \varphi_2 \equiv 0$) and
- stationary if the shifted process $w(\cdot - \mathbf{r}_0)$ has the same finite-dimensional marginals as w .

The properties in Definition 2 can be inferred from the characteristic functional of the process. Specifically, the independence property corresponds to the condition

$$\widehat{\mathcal{P}}_w(\varphi_1 + \varphi_2) = \widehat{\mathcal{P}}_w(\varphi_1) \widehat{\mathcal{P}}_w(\varphi_2) \quad (7)$$

whenever φ_1 and φ_2 have disjoint supports [GV64]. Moreover, w is stationary iff. it has the same characteristic functional as the process $w(\cdot - \mathbf{r}_0)$ defined by $\langle w(\cdot - \mathbf{r}_0), \varphi \rangle = \langle w, \varphi(\cdot + \mathbf{r}_0) \rangle$, i.e. iff. $\forall \varphi \in \mathcal{D}$ and $\mathbf{r}_0 \in \mathbb{R}^d$,

$$\widehat{\mathcal{P}}_w(\varphi(\cdot - \mathbf{r}_0)) = \widehat{\mathcal{P}}_w(\varphi). \quad (8)$$

The functional

$$\widehat{\mathcal{P}}(\varphi) = \exp \left(\int_{\mathbb{R}^d} f(\varphi(\mathbf{r})) d\mathbf{r} \right) \quad (9)$$

with $f(0) = 0$ satisfies the equations (7) and (8). Moreover, Gelfand and Vilenkin give necessary and sufficient conditions on f so that the functional is continuous and positive-definite over \mathcal{D} , and hence, defines a valid innovation process w .

Theorem 2 (Gelfand-Vilenkin). Define $\widehat{\mathcal{P}}(\varphi) = \exp \left(\int_{\mathbb{R}^d} f(\varphi(\mathbf{r})) d\mathbf{r} \right)$ on \mathcal{D} where f is a continuous function from \mathbb{R} to \mathbb{C} with $f(0) = 0$. The following conditions are equivalent:

- (i) There exists a (unique) probability measure \mathcal{P}_w on \mathcal{D}' such that

$$\widehat{\mathcal{P}}(\varphi) = \widehat{\mathcal{P}}_w(\varphi).$$

- (ii) The functional $\widehat{\mathcal{P}}$ is a continuous, positive-definite, and normalized ($\widehat{\mathcal{P}}(0) = 1$) functional on \mathcal{D} .

- (iii) There exist $\mu \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}^+$, and a Lévy measure V with $\int_{\mathbb{R} \setminus \{0\}} \min(1, a^2) V(da) < \infty$ such that

$$f(\omega) = j\mu\omega - \frac{\sigma^2\omega^2}{2} + \int_{\mathbb{R} \setminus \{0\}} (e^{j\omega a} - 1 - j\omega a \mathbb{1}_{|a| < 1}) V(da). \quad (10)$$

A function f that can be represented in the form of (10) is called a *Lévy exponent*. The function is alternatively characterized by the triplet (μ, σ^2, V) known as the *Lévy triplet*.

Definition 3. A white Lévy noise, or an innovation process on \mathcal{D}' is a generalized stochastic process w with probability measure \mathcal{P}_w on \mathcal{D}' that is characterized by $\widehat{\mathcal{P}}_w(\varphi) = \exp \left(\int_{\mathbb{R}^d} f(\varphi(\mathbf{r})) d\mathbf{r} \right)$ for some Lévy exponent f . In addition, the functional $F(\varphi) = \log \widehat{\mathcal{P}}_w(\varphi)$ is called the generalized Lévy exponent associated with w .

A white Lévy noise on \mathcal{D}' has an independent value at every point and is stationary, which justifies the “white noise” nomenclature. By Theorem 2, we have a one-to-one correspondence between Lévy exponents f and the white Lévy noises on \mathcal{D}' . Interestingly, one has the same one-to-one correspondence between the family of infinite-divisible probability laws and Lévy exponents. Indeed, p_X is an infinite-divisible pdf if and only if $\widehat{p}_X(\omega) = e^{f(\omega)}$ where $f(\cdot)$ is a valid Lévy exponent [Sat94].

Gelfand and Vilenkin’s constructive result on the characteristic functional of an innovation process on \mathcal{D} resolves the central barrier of the positive-definiteness requirement in applying the Minlos-Bochner theorem. Indeed, we shall show in Proposition 2 that, for extending Theorem 2 to larger spaces of test functions, we only require to prove the continuity of the functional (9) as the positive-definiteness is automatically inherited.

Proposition 2. *Let \mathcal{T} be any of \mathcal{S} , L^p , or $L^p \cap L^q$ for $p, q > 0$. Assume f is a Lévy exponent such that the functional $\widehat{\mathcal{P}}(\varphi) = \exp(\int_{\mathbb{R}^d} f(\varphi(\mathbf{r})) d\mathbf{r})$ is well-defined (namely, $f(\varphi(\mathbf{r})) \in L^1$) and is continuous by the natural topology of \mathcal{T} . Then, $\widehat{\mathcal{P}}$ is also positive-definite over \mathcal{T} .*

Note that the topological structure of $L^p \cap L^q$ depends on the relative values of p and q with respect to 1. If, for instance, $p < 1 \leq q$, then $L^p \cap L^q$ is a metric space with $d_p(f, g) + \|f - g\|_q$ (see Table 1 or [Rud91]).

Proof. From Theorem 2 we know that $\widehat{\mathcal{P}}$ is well-defined, continuous, normalized and positive-definite over \mathcal{D} . We then use a density argument to extend the positive-definiteness to \mathcal{T} . Indeed, \mathcal{D} is dense in all possible \mathcal{T} of Proposition 2. This result is well-known for \mathcal{S} and the L^p spaces with $p \geq 1$. The proof for L^p spaces with $0 < p < 1$ and for $L^p \cap L^q$ spaces is also similar.

Let $\varphi_1, \dots, \varphi_N \in \mathcal{T}$ and $a_1, \dots, a_N \in \mathbb{C}$. Since \mathcal{D} is dense in \mathcal{T} , there exist sequences $(\varphi_n^k)_{1 \leq n \leq N, k \in \mathbb{N}}$ of functions in \mathcal{D} such that $\lim_{k \rightarrow +\infty} \varphi_n^k = \varphi_n$ for all n . Then, by using the continuity of $\widehat{\mathcal{P}}$ over \mathcal{T} , we obtain that

$$\sum_{1 \leq i, j \leq N} a_i \bar{a}_j \widehat{\mathcal{P}}(\varphi_i - \varphi_j) = \lim_{k \rightarrow +\infty} \left(\sum_{1 \leq i, j \leq N} a_i \bar{a}_j \widehat{\mathcal{P}}(\varphi_i^k - \varphi_j^k) \right) \geq 0. \quad (11)$$

□

Innovation Processes on \mathcal{S}'

We recall that the Minlos-Bochner theorem is valid for any nuclear space including $\mathcal{S}' \subset \mathcal{D}'$, which allows us to generalize Definition 2 to \mathcal{S}' . Moreover, it is possible to characterize the independence and the stationarity of a generalized process on \mathcal{S}' directly on its characteristic functional in the same way we did for \mathcal{D}' in (7) and (8). Next, we introduce a sufficient condition on the Lévy exponent f (more precisely, on the Lévy measure) to extend the notion of innovation process to \mathcal{S}' by applying Theorem 1. We first give some definitions.

Definition 4. *Let $\mathcal{M}(\mathbb{R} \setminus \{0\})$ be the set of Radon measures on $\mathbb{R} \setminus \{0\}$. For $V \in \mathcal{M}(\mathbb{R} \setminus \{0\})$ and $k \geq 0$, we denote*

$$\mu_k(V) = \int_{\mathbb{R} \setminus \{0\}} |a|^k V(da), \quad (12)$$

$$\mu_k^0(V) = \int_{0 < |a| < 1} |a|^k V(da), \quad (13)$$

$$\mu_k^\infty(V) = \int_{|a| \geq 1} |a|^k V(da). \quad (14)$$

with $\mu_k(V) = \mu_k^0(V) + \mu_k^\infty(V)$. Further, we define

$$\mathcal{M}(p, q) = \left\{ V \in \mathcal{M}(\mathbb{R} \setminus \{0\}) \mid \mu_q^0 < \infty \text{ and } \mu_p^\infty < \infty \right\}. \quad (15)$$

Hence, the set of Lévy measures corresponds to $\mathcal{M}(0, 2)$. We also define the set of Lévy-Schwartz measures as a subset of Lévy measures adapted for extending the framework of Gelfand and Vilenkin to the Schwartz space \mathcal{S} (see Theorem 3) by

$$\mathcal{M}(0^+, 2) = \bigcup_{\epsilon > 0} \mathcal{M}(\epsilon, 2) \subset \mathcal{M}(0, 2). \quad (16)$$

It is not difficult to check the following properties of the sets $\mathcal{M}(p, q)$:

- $\mathcal{M}(p_1, q_1) \cap \mathcal{M}(p_2, q_2) = \mathcal{M}(\max(p_1, p_2), \min(q_1, q_2))$,
- $\mathcal{M}(p_1, q_1) \cup \mathcal{M}(p_2, q_2) = \mathcal{M}(\min(p_1, p_2), \max(q_1, q_2))$,
- $\mathcal{M}(p_1, q_1) \subset \mathcal{M}(p_2, q_2) \Leftrightarrow p_1 \geq p_2 \text{ and } q_1 \leq q_2$.

The interest of Definition 4 is to focus separately on the behavior of V around 0 and at infinities. It also helps in classifying the innovation processes according to their Lévy measure. For instance, Poisson innovations correspond to $V \in \mathcal{M}(0, 0)$ while innovations with finite variance are obtained for $V \in \mathcal{M}(2, 2)$. In Theorem 3, we state our main result concerning innovation processes over \mathcal{S}' .

Theorem 3 (Tempered innovation processes). *Suppose that f is a Lévy exponent with triplet (μ, σ^2, V) , where V is a Lévy-Schwartz measure (Definition 4). Then, there exists a unique measure \mathcal{P} on \mathcal{S}' such that*

$$\widehat{\mathcal{P}}(\varphi) = \int_{\mathcal{S}'} e^{j\langle u, \varphi \rangle} d\mathcal{P}(u) = \exp \left(\int_{\mathbb{R}^d} f(\varphi(\mathbf{r})) d\mathbf{r} \right), \quad \forall \varphi \in \mathcal{S}. \quad (17)$$

The underlying generalized stochastic process w associated with \mathcal{P} is called a tempered innovation process or a white Lévy-Schwartz noise.

Proof. The function space \mathcal{S} is nuclear, which justifies the application of the Minlos-Bochner theorem. Obviously, $\widehat{\mathcal{P}}(0) = 1$, and $\widehat{\mathcal{P}}(\varphi)$ is also positive-definite, given it is continuous (Proposition 2). However, note that it is not *a priori* evident that $f(\varphi(\mathbf{r}))$ is even integrable for $\varphi \in \mathcal{S}$, whereas the integrability is easily understood for $\varphi \in \mathcal{D}$, since f is continuous and φ is of finite support. We prove Theorem 3 by successively establishing the integrability of $f(\varphi(\mathbf{r}))$ for $\varphi \in \mathcal{S}$ and the continuity of the functional $\widehat{\mathcal{P}}$ on \mathcal{S} .

The proof is based on a control on the generalized Lévy exponent developed in Section 3.3. To use this result we first remark that \mathcal{S} is a subspace of all L^p spaces. Moreover, the continuity of a functional over \mathcal{S} implies its continuity over any L^p space with $p > 0$.

Since V is a Lévy-Schwartz measure, there exists $0 < \epsilon \leq 1$ such that $V \in \mathcal{M}(\epsilon, 2)$. Using Corollary 1 (Section 3.3), we know that there exist κ_1 and $\kappa_2 \geq 0$ such that, for all $\varphi \in \mathcal{S}$,

$$\int_{\mathbb{R}^d} |f(\varphi(\mathbf{r}))| d\mathbf{r} \leq \mu \|\varphi\|_1 + \frac{\sigma^2}{2} \|\varphi\|_2^2 + \kappa_1 \|\varphi\|_\epsilon^\epsilon + \kappa_2 \|\varphi\|_2^2.$$

As $\|\varphi\|_p$ is finite for all $p > 0$, we conclude that $F(\varphi)$ is well-defined over \mathcal{S} . In addition, from Proposition 4 (Section 3.3) we know that there exist ν_1 and $\nu_2 \geq 0$ such that, for all $\varphi, \psi \in \mathcal{S}$,

$$|F(\varphi) - F(\psi)| \leq \nu_1 \sqrt{(\|\varphi\|_\epsilon^\epsilon + \|\psi\|_\epsilon^\epsilon)(\|\varphi - \psi\|_\epsilon^\epsilon)} + \nu_2 \sqrt{(\|\varphi\|_2^2 + \|\psi\|_2^2)(\|\varphi - \psi\|_2^2)}.$$

Consequently, if $\varphi_n \rightarrow \varphi$ in \mathcal{S} , then $F(\varphi_n) \rightarrow F(\varphi)$ in \mathbb{C} . This shows that $\widehat{\mathcal{P}}(\varphi) = \exp(F(\varphi))$ is continuous over \mathcal{S} , which completes the proof by applying the Minlos-Bochner theorem. \square

The restriction $V \in \mathcal{M}(0^+, 2)$ in Theorem 3 is extremely mild and plays no role in all cases of practical interest (Table 2). Yet, it is possible to construct examples of Lévy measures $V \in \mathcal{M}(0, 2) \setminus \mathcal{M}(0^+, 2)$ such as

$$V(da) = \frac{da}{|a| \log^2(2 + |a|)}. \quad (18)$$

We give in Table 2 the main examples of white Lévy-Schwartz noises: Gaussian noises, symmetric α -stable noises (noted by S α S, see [ST94]), Variance Gamma noise (which includes the Laplace distribution and is linked with TV-regularization [BKNU13]), and Poisson noises.

Table 2: Tempered innovation processes

Distribution and parameters	Lévy triplet $(\mu, \sigma^2, V(da))$	Generalized Lévy exponent $F(\varphi)$
Gaussian $(\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+$	$(\mu, \sigma^2, 0)$	$j\mu(\int \varphi) - \frac{\sigma^2 \ \varphi\ _2^2}{2}$
SaS $(\alpha, \gamma) \in (0, 2) \times \mathbb{R}_+$	$\left(0, 0, \frac{C_{\alpha, \gamma}}{ a ^{\alpha+1}} da\right)$ $V \in \mathcal{M}(\alpha^-, \alpha^+)$	$-\gamma^\alpha \ \varphi\ _\alpha^\alpha$
Variance Gamma $\lambda \in \mathbb{R}_+$	$\left(0, 0, \frac{e^{-\lambda a }}{ a } da\right)$ $V \in \mathcal{M}(\infty, 0^+)$	$\int_{\mathbb{R}^d} \log\left(\frac{\lambda^2}{\lambda^2 + \varphi(\mathbf{r})^2}\right) d\mathbf{r}$
Poisson $\lambda > 0, P$ probability measure	$(0, 0, \lambda P(da))$ $V \in \mathcal{M}(0^+, 0)$	$-j\lambda\mu_1^0(P)(\int \varphi) + \lambda \int_{\mathbb{R}^d} \int_{\mathbb{R} \setminus \{0\}} (e^{ia\varphi(\mathbf{r})} - 1) P(da) d\mathbf{r}$

Link Between Innovation Processes on \mathcal{D}' and \mathcal{S}'

Let f be a Lévy exponent with a Lévy-Schwartz measure. According to Theorems 2 and 3, we can define,

- a measure $\mathcal{P}_{\mathcal{D}'}$ on \mathcal{D}' such that $\widehat{\mathcal{P}_{\mathcal{D}'}}(\varphi) = \exp\left(\int f(\varphi(\mathbf{r})) d\mathbf{r}\right)$ for $\varphi \in \mathcal{D}$, and
- a measure $\mathcal{P}_{\mathcal{S}'}$ on \mathcal{S}' such that $\widehat{\mathcal{P}_{\mathcal{S}'}}(\varphi) = \exp\left(\int f(\varphi(\mathbf{r})) d\mathbf{r}\right)$ for $\varphi \in \mathcal{S}$, respectively.

We discuss here the compatibility of the two measures. Let $\mathcal{N} = \mathcal{D}$ or \mathcal{S} . First, we recall the method of constructing a measure on the nuclear space \mathcal{N} . For given $\varphi = (\varphi_1, \dots, \varphi_N) \in \mathcal{N}^N$ and a Borelian subset B of \mathbb{R}^N , a cylindric set is defined as

$$A_{\varphi, B}^{\mathcal{N}'} = \{u \in \mathcal{N}', (\langle u, \varphi_1 \rangle, \dots, \langle u, \varphi_N \rangle) \in B\}. \quad (19)$$

If $\mathcal{C}_{\mathcal{N}'}$ denotes the collection of all such cylindric sets, then, according to the Minlos-Bochner theorem, the σ -field $\mathcal{A}_{\mathcal{N}'} = \sigma(\mathcal{C}_{\mathcal{N}'})$ generated by the cylindric sets properly specifies a probability measure on \mathcal{N}' . In Proposition 3, we compare the σ -fields $\mathcal{A}_{\mathcal{S}'}$ and $\mathcal{A}_{\mathcal{D}'}$. Note that it is not obvious *a priori* that the two σ -fields are closely related. The main difficulty is to see that the space \mathcal{S}' itself is an element of $\mathcal{A}_{\mathcal{D}'}$. The result of Proposition 3 is necessary to be able to compare the two measures $\mathcal{P}_{\mathcal{S}'}$ and $\mathcal{P}_{\mathcal{D}'}$.

Proposition 3. *We have the relations*

$$\mathcal{A}_{\mathcal{S}'} = \{A \cap \mathcal{S}' \mid A \in \mathcal{A}_{\mathcal{D}'}\} \quad (20)$$

$$\subset \mathcal{A}_{\mathcal{D}'}. \quad (21)$$

Proof. We decompose the proof in four steps.

(1) We denote $\mathcal{C}_{\mathcal{N}'}^\Omega$ the collection of cylindric set $A_{\varphi, B}^{\mathcal{N}'}$ with Ω an open set of \mathbb{R}^N . We claim that $\sigma(\mathcal{C}_{\mathcal{N}'}^\Omega) = \mathcal{A}_{\mathcal{N}'}$. This result is obtained by a transfinite induction using the fact that the open sets generates the σ -field of the Borelian sets.

(2) We show that $\mathcal{S}' \in \mathcal{A}_{\mathcal{D}'}$. For $\alpha \in \mathbb{N}$, we consider (see Table 1) $N_\alpha(\varphi) = \sum_{0 \leq |\mathbf{k}| \leq \alpha} \|\partial^{\mathbf{k}} \varphi\|_{\infty, \alpha}$ where

$$|\mathbf{k}| = k_1 + \dots + k_d \text{ and } \partial^{\mathbf{k}} \varphi = \frac{\partial^{\mathbf{k}}}{\partial r_1^{k_1} \dots \partial r_d^{k_d}} \varphi.$$

A generalized function $u \in \mathcal{D}'$ is tempered iff. there exist $\alpha \in \mathbb{N}$ and $C > 0$ such that, for all $\varphi \in \mathcal{D}$, $|\langle u, \varphi \rangle| \leq C N_\alpha(\varphi)$. Then, u can be uniquely extended to a continuous linear form on \mathcal{S} . The space

\mathcal{S}' is identified as a subspace of \mathcal{D}' [Bon01]. In addition, we know that \mathcal{S} is separable: there exists a sequence $(\varphi_n) \in \mathcal{S}^{\mathbb{N}}$ that is dense in \mathcal{S} . Because \mathcal{D} is dense in \mathcal{S} , we can also impose that the φ_n are in \mathcal{D} . Consequently, we have

$$\mathcal{S}' = \bigcup_{C \in \mathbb{N}} \bigcup_{\alpha \in \mathbb{N}} \bigcap_{n \in \mathbb{N}} A_{\varphi_n, [-C N_{\alpha} \varphi_n, C N_{\alpha} \varphi_n]} \in \mathcal{A}_{\mathcal{D}'}. \quad (22)$$

A direct consequence is that $\{A \cap \mathcal{S}' \mid A \in \mathcal{A}_{\mathcal{D}'}\} \subset \mathcal{A}_{\mathcal{D}'}$.

(3) First, we remark that $\{A \cap \mathcal{S}' \mid A \in \mathcal{A}_{\mathcal{D}'}\} \cap \mathcal{S}'$ is a σ -field on \mathcal{S}' (as a restriction of a σ -field on \mathcal{D}') containing $\mathcal{C}_{\mathcal{D}'}^{\Omega} \cap \mathcal{S}'$ and then $\sigma(\mathcal{C}_{\mathcal{D}'}^{\Omega} \cap \mathcal{S}')$. Consequently, it is enough to show that $\mathcal{C}_{\mathcal{S}'}^{\Omega} \subset \sigma(\mathcal{C}_{\mathcal{D}'}^{\Omega} \cap \mathcal{S}')$. Let us fix $\psi_1, \dots, \psi_N \in \mathcal{S}$ and Ω an open set of \mathbb{R}^N . Let $(\varphi_{n,k})_{n=1, \dots, N, k \in \mathbb{N}}$ be N sequences of functions in \mathcal{D} converging in \mathcal{S} to ψ_n for all $n \in \{1, \dots, N\}$. Because Ω is open, for all $u \in \mathcal{S}'$, $(\langle u, \psi_1 \rangle, \dots, \langle u, \psi_N \rangle) \in \Omega$ iff. $(\langle u, \varphi_{1,k} \rangle, \dots, \langle u, \varphi_{N,k} \rangle) \in \Omega$ for k large enough. Moreover, because $\mathcal{D} \subset \mathcal{S}$, we have $A_{\psi, \Omega}^{\mathcal{D}'} \cup \mathcal{S}' = A_{\psi, \Omega}^{\mathcal{S}'}$. Thus,

$$\begin{aligned} A_{\psi, B}^{\mathcal{S}'} &= \bigcup_{p \in \mathbb{N}} \bigcap_{k \geq p} (A_{\varphi_k, B}^{\mathcal{S}'}) \\ &= \bigcup_{p \in \mathbb{N}} \bigcap_{k \geq p} (A_{\varphi_k, B}^{\mathcal{D}'} \cap \mathcal{S}') \\ &\in \sigma(\mathcal{C}_{\mathcal{D}'}^{\Omega} \cap \mathcal{S}'). \end{aligned}$$

As a consequence, $\mathcal{A}_{\mathcal{S}'} \subset \{A \cap \mathcal{S}' \mid A \in \mathcal{A}_{\mathcal{D}'}\}$.

(4) For the other inclusion, we first notice that $\{A \cap \mathcal{S}' \mid A \in \mathcal{A}_{\mathcal{D}'}\} = \sigma(\mathcal{C}_{\mathcal{D}'}^{\Omega} \cap \mathcal{S}')$ (the restriction of the generator family generates the restrictive σ -field). Thus, we just need to prove that $A_{\psi, \Omega}^{\mathcal{D}'} \cup \mathcal{S}' \in \mathcal{A}_{\mathcal{S}'}$ for all $\psi \in \mathcal{D}^N$ and Ω an open set of \mathbb{R}^N , which is obvious because, as we said, $A_{\psi, \Omega}^{\mathcal{D}'} \cup \mathcal{S}' = A_{\psi, \Omega}^{\mathcal{S}'}$. Consequently, $\{A \cap \mathcal{S}' \mid A \in \mathcal{A}_{\mathcal{D}'}\} \subset \mathcal{A}_{\mathcal{S}'}$. With (3), we obtain that $\{A \cap \mathcal{S}' \mid A \in \mathcal{A}_{\mathcal{D}'}\} = \mathcal{A}_{\mathcal{S}'}$. \square

We now focus on measures over $\mathcal{A}_{\mathcal{D}'}$ and $\mathcal{A}_{\mathcal{S}'}$ that define innovation processes.

Theorem 4. *Let f be a Lévy exponent with Lévy measure $V \in \mathcal{M}(0^+, 2)$. By $\mathcal{P}_{\mathcal{D}'}$ and $\mathcal{P}_{\mathcal{S}'}$ we denote the measures on \mathcal{D}' and \mathcal{S}' , respectively, that are defined by the characteristic functional $\exp(\int_{\mathbb{R}^d} f(\varphi(\mathbf{r})) d\mathbf{r})$ (over \mathcal{D} and \mathcal{S} , respectively). The two measures are compatible in the sense that*

$$\forall A \in \mathcal{A}_{\mathcal{S}'}, \mathcal{P}_{\mathcal{D}'}(A) = \mathcal{P}_{\mathcal{S}'}(A). \quad (23)$$

In particular, $\mathcal{P}_{\mathcal{D}'}(\mathcal{S}') = 1$ and $\mathcal{P}_{\mathcal{D}'}(\mathcal{D}' \setminus \mathcal{S}') = 0$.

Proof. From $\mathcal{P}_{\mathcal{S}'}$, we define a new measure on \mathcal{D}' by $\mathcal{P}(A) = \mathcal{P}_{\mathcal{S}'}(A \cap \mathcal{S}')$ for $A \in \mathcal{A}_{\mathcal{D}'}$. We claim that $\mathcal{P} = \mathcal{P}_{\mathcal{D}'}$. For $\varphi \in \mathcal{D}$, we have

$$\begin{aligned} \widehat{\mathcal{P}}(\varphi) &= \int_{\mathcal{D}'} e^{j\langle u, \varphi \rangle} d\mathcal{P}(u) \\ &= \int_{\mathcal{S}'} e^{j\langle u, \varphi \rangle} d\mathcal{P}(u) \\ &= \int_{\mathcal{S}'} e^{j\langle u, \varphi \rangle} d\mathcal{P}_{\mathcal{S}'}(u) \end{aligned} \quad (24)$$

$$\begin{aligned} &= \exp\left(\int_{\mathbb{R}^d} f(\varphi(\mathbf{r})) d\mathbf{r}\right) \\ &= \widehat{\mathcal{P}_{\mathcal{D}'}}(\varphi). \end{aligned} \quad (25)$$

We used that $\mathcal{P}(\mathcal{D}' \setminus \mathcal{S}') = 0$ in (24) and that \mathcal{P} restricted to \mathcal{S}' coincides with $\mathcal{P}_{\mathcal{S}'}$ in (25). The Minlos-Bochner theorem ensures that $\mathcal{P} = \mathcal{P}_{\mathcal{D}'}$. Fix $A \in \mathcal{S}'$. According to Proposition 3, $A \in \mathcal{A}_{\mathcal{D}'}$ and $\mathcal{P}_{\mathcal{D}'}(A)$ is well-defined. Consequently, we have that $\mathcal{P}_{\mathcal{D}'}(A) = \mathcal{P}(A) = \mathcal{P}_{\mathcal{S}'}(A \cap \mathcal{S}') = \mathcal{P}_{\mathcal{S}'}(A)$. For $A = \mathcal{S}'$, we obtain $\mathcal{P}_{\mathcal{D}'}(\mathcal{S}') = 1$ and, consequently, $\mathcal{P}_{\mathcal{D}'}(\mathcal{D}' \setminus \mathcal{S}') = 0$. \square

The essential fact is that the theory of Gelfand already defines probability measures concentrated on the tempered generalized functions.

3.2 Sparse Processes

In the remainder of the paper, we restrict our attention to $\mathcal{N} = \mathcal{S}$. In Section 3.1, we defined stochastic processes on \mathcal{S}' , which are *bona fide* innovation processes. According to [UTS11], these innovation processes are split into two categories: (i) white Gaussian noises corresponding to zero Lévy measure and the Lévy exponent $f(\omega) = j\mu\omega - \frac{\sigma^2}{2}\omega^2$, (ii) non-Gaussian white noises with non zero Lévy measures, which are referred to as sparse innovation processes. The reason is that all non-Gaussian infinite-divisible distributions are necessarily more compressible than Gaussians [UTS11].

Our next goal is to define processes s such that $Ls = w$ is an innovation process. With the rationale as above, the processes leading to non-Gaussian innovations w will be called sparse. The fact that the linear operator L whitens s implies that there exists a deterministic operator L^{-1} that induces the dependency structure of the process s . If we follow the formal equalities $\langle s, \varphi \rangle = \langle L^{-1}w, \varphi \rangle = \langle w, L^{*-1}\varphi \rangle$, we then interpret the model $Ls = w$ as

$$\widehat{\mathcal{P}}_s(\varphi) = \exp \left(\int_{\mathbb{R}^d} f(L^{*-1}\varphi(\mathbf{r})) d\mathbf{r} \right). \quad (26)$$

However, we do not know *a priori* if (26) defines a valid characteristic functional. This is especially true if (a) the operator L^{*-1} is continuous from \mathcal{S} to some function space \mathcal{T} (not necessarily nuclear) and if (b) the functional $\psi \mapsto \exp \left(\int_{\mathbb{R}^d} f(\psi(\mathbf{r})) d\mathbf{r} \right)$ is well-defined and continuous over \mathcal{T} . More precisely, we are examining the compatibility of the Lévy exponent f with the linear operator L to define a characteristic functional. Concretely, we are concerned with the function spaces $\mathcal{T} = \mathcal{S}, \mathcal{R}, L^p$ -spaces, and intersections of L^p -spaces. We state in Theorem 5 a sufficient compatibility condition. For the sake of generality, the operator L^{*-1} is replaced with the generic operator T .

Theorem 5. (*Compatibility conditions*) Let f be a Lévy exponent with triplet (μ, σ^2, V) and let T be a linear operator from \mathcal{S} to \mathcal{S}' . Suppose we have $0 < p_{\min} \leq p_{\max} \leq 2$ and

- $V \in \mathcal{M}(p_{\min}, p_{\max})$,
- $p_{\min} \leq 1$, if $\mu \neq 0$ or V non-symmetric,
- $p_{\max} = 2$, if $\sigma^2 \neq 0$, and
- T is a linear and continuous operator from \mathcal{S} to $L^{p_{\min}} \cap L^{p_{\max}}$.

Then, there exists a unique probability measure \mathcal{P}_s on \mathcal{S}' with

$$\widehat{\mathcal{P}}_s(\varphi) = \exp \left(\int_{\mathbb{R}^d} f(T\varphi(\mathbf{r})) d\mathbf{r} \right). \quad (27)$$

Proof. We apply the Minlos-Bochner theorem to the functional $\widehat{\mathcal{P}}(\varphi) = \exp \left(\int_{\mathbb{R}^d} f(T\varphi(\mathbf{r})) d\mathbf{r} \right)$. It is normalized because the linearity of T implies $f(T\{0\}) = f(0) = 0$. The linearity of T also enables us to conclude the positive-definiteness of the functional from Proposition 2, given that it is well-defined and continuous over \mathcal{S} . The continuity is established by applying the bounds in Section 3.3 on the generalized Lévy exponent $F(\varphi) = \int_{\mathbb{R}^d} f(\varphi(\mathbf{r})) d\mathbf{r}$. In particular, Corollary 1 implies the existence of κ_1 and $\kappa_2 \geq 0$ such that, for all $\varphi \in \mathcal{S}$,

$$\int_{\mathbb{R}^d} |f(T\varphi(\mathbf{r}))| d\mathbf{r} \leq \mu \|T\varphi\|_1 + \frac{\sigma^2}{2} \|T\varphi\|_2^2 + \kappa_1 \|T\varphi\|_{p_{\min}}^{p_{\min}} + \kappa_2 \|T\varphi\|_{p_{\max}}^{p_{\max}},$$

and the assumptions on f and T ensure that the integral is finite. Thus, $\widehat{\mathcal{P}}$ is well-defined. Under the assumptions of Theorem 5, we can also apply Proposition 4 and find ν_1 and $\nu_2 \geq 0$ such that, for all $\varphi, \psi \in \mathcal{S}$,

$$\begin{aligned} |F(T\varphi) - F(T\psi)| &\leq \nu_1 \sqrt{(\|T\varphi\|_{p_{\min}}^{p_{\min}} + \|T\psi\|_{p_{\min}}^{p_{\min}}) \|T\varphi - T\psi\|_{p_{\min}}^{p_{\min}}} \\ &\quad + \nu_2 \sqrt{(\|T\varphi\|_{p_{\max}}^{p_{\max}} + \|T\psi\|_{p_{\max}}^{p_{\max}}) \|T\varphi - T\psi\|_{p_{\max}}^{p_{\max}}}. \end{aligned}$$

Now, if $\varphi_n \rightarrow \varphi$ in \mathcal{S} , then, $T\varphi_n \rightarrow T\varphi$ in $L^{p_{\min}} \cap L^{p_{\max}}$ (continuity of T) and $F(T\varphi_n) \rightarrow F(T\varphi)$ in \mathbb{C} . Hence, $\widehat{\mathcal{P}}(\varphi)$ is continuous over \mathcal{S} . According to Theorem 1, there exists a unique s such that $\widehat{\mathcal{P}}_s = \widehat{\mathcal{P}}$. □

Comparison with p -admissibility. Theorem 5 gives a compatibility condition between f and L . Another condition, called p -admissibility, was introduced in [UTS11]. A Lévy exponent f is said to be p -admissible if $|f(\omega)| + |\omega f'(\omega)| \leq C|\omega|^p$ for all $\omega \in \mathbb{R}$, where $1 \leq p < +\infty$ and C is a positive constant. Although p -admissibility is sufficient in many practical cases, we argue that it is generally more restrictive than the assumptions in Theorem 5.

- The p -admissibility condition is restricted to $p \geq 1$ and requires the differentiability of the Lévy exponent. The most natural sufficient condition to assure differentiability is that $\mu_1(V) < +\infty$ (or $\mu_1^\infty(V) < +\infty$ when V is symmetric). In contrast, Theorem 5 does not impose the differentiability constraint and includes scenarios with $p < 1$.
- The notion of $\mathcal{M}(p, q)$ introduced in Definition 4 distinguishes the limitations imposed by the Lévy measure V at $a \rightarrow 0$ and $a \rightarrow \infty$. As a result, Theorem 5 allows for a richer family of Lévy exponents f . For instance, suppose that $f = f_\alpha + f_\beta$ is the sum of two S α S Lévy exponents with $\alpha < \beta$. Then, although f_α and f_β can be α -admissible and β -admissible, respectively, f is not p -admissible for any $p > 0$. It is not hard to check that f is covered by Theorem 5.
- The assumptions of Theorem 5 can also be slightly restrictive. The S α S case is a generic example. We denote V_α the Lévy measure of the S α S Lévy exponent f_α . Then, because $\mu_\alpha^\infty(V_\alpha) = \mu_\alpha^0(V_\alpha) = +\infty$, the Theorem 5 only allows for $V_\alpha \in \mathcal{M}(\alpha^-, \alpha^+) = \bigcup_{\epsilon > 0} \mathcal{M}(\alpha - \epsilon, \alpha + \epsilon)$, but the condition $\varphi \in L^\alpha$ is clearly sufficient (and necessary) in practice. However, we know that $f_\alpha(\omega) = -|\omega|^\alpha$ is α -admissible.

3.3 Continuity of Characteristic Functionals

This section is devoted to the derivation of bounds on the generalized Lévy exponents to conclude the continuity results required in Theorems 3 and 5. We first introduce some notations and useful inequalities.

Definition 5. Let $p > 0$, $x, y \in \mathbb{R}$, and $f, g \in L^p(\mathbb{R}^d)$. We define

$$h_p(x, y) = \sqrt{(|x|^p + |y|^p) |x - y|^p}, \quad (28)$$

$$H_p(f, g) = \sqrt{(\|f\|_p^p + \|g\|_p^p) \|f - g\|_p^p}. \quad (29)$$

Lemma 1. Let $p, q > 0$.

(i) For all $x, y \in \mathbb{R}$ we have that

$$|x - y|^p \leq \max\left(1, 2^{\frac{p-1}{2}}\right) h_p(x, y), \quad (30)$$

$$|x^2 - y^2|^{p/2} \leq \max\left(1, 2^{\frac{p-1}{2}}\right) h_p(x, y). \quad (31)$$

(ii) For $f, g \in L^p(\mathbb{R}^d)$, we have that

$$\int_{\mathbb{R}^d} h_p(f(\mathbf{r}), g(\mathbf{r})) d\mathbf{r} \leq H_p(f, g). \quad (32)$$

(iii) For $f, g \in L^p \cap L^q$ and $\lambda \in [0, 1]$,

$$H_{\lambda p + (1-\lambda)q}(f, g) \leq \sqrt{\lambda} H_p(f, g) + \sqrt{1-\lambda} H_q(f, g). \quad (33)$$

Proof. For $p \geq 1$, it follows from Jensen's inequality that $|x \pm y|^p \leq 2^{p-1}(|x|^p + |y|^p)$. Moreover, for $0 < p < 1$, we have that $|x \pm y|^p \leq |x|^p + |y|^p$. Consequently,

$$\begin{aligned} |x - y|^p &= \sqrt{|x - y|^p} \sqrt{|x - y|^p} \\ &\leq \max(1, 2^{\frac{p-1}{2}}) h_p(x, y) \end{aligned}$$

and

$$\begin{aligned} |x^2 - y^2|^{p/2} &= \sqrt{|x+y|^p} \sqrt{|x-y|^p} \\ &\leq \max(1, 2^{\frac{p-1}{2}}) h_p(x, y). \end{aligned}$$

Let now $f, g \in L^p(\mathbb{R}^d)$. By invoking the Cauchy-Schwartz inequality, we can verify that

$$\begin{aligned} \int_{\mathbb{R}^d} h_p(f(\mathbf{r}), g(\mathbf{r})) d\mathbf{r} &= \int_{\mathbb{R}^d} \sqrt{|f(\mathbf{r})|^p + |g(\mathbf{r})|^p} \sqrt{|f(\mathbf{r}) - g(\mathbf{r})|^p} d\mathbf{r} \\ &\leq \sqrt{\int_{\mathbb{R}^d} (|f(\mathbf{r})|^p + |g(\mathbf{r})|^p) d\mathbf{r}} \sqrt{\int_{\mathbb{R}^d} |f(\mathbf{r}) - g(\mathbf{r})|^p d\mathbf{r}} \\ &= H_p(f, g). \end{aligned}$$

To prove (iii), we define $F(\mathbf{r}_1, \mathbf{r}_2) = f(\mathbf{r}_1)(g - f)(\mathbf{r}_2)$ and $G(\mathbf{r}_1, \mathbf{r}_2) = g(\mathbf{r}_1)(f - g)(\mathbf{r}_2)$. As a consequence, $H_p(f, g) = \sqrt{\|F\|_p^p + \|G\|_p^p}$. We now write that

$$\begin{aligned} H_{\lambda p + (1-\lambda)q}(f, g) &= \sqrt{\|F\|_{\lambda p + (1-\lambda)q}^{\lambda p + (1-\lambda)q} + \|G\|_{\lambda p + (1-\lambda)q}^{\lambda p + (1-\lambda)q}} \\ &\leq \sqrt{\lambda \|F\|_p^p + (1-\lambda) \|F\|_q^q + \lambda \|G\|_p^p + (1-\lambda) \|G\|_q^q} \\ &\quad [\text{using the convexity of } p \mapsto a^p \text{ for } a \geq 0] \\ &\leq \sqrt{\lambda (\|F\|_p^p + \|G\|_p^p) + (1-\lambda) (\|F\|_q^q + \|G\|_q^q)} \\ &\quad [\text{using the concavity of } \sqrt{\cdot}] \\ &= \sqrt{\lambda} H_p(f, g) + \sqrt{1-\lambda} H_q(f, g). \end{aligned}$$

□

The key step towards obtaining continuity bounds is to control the non-Gaussian part g of the Lévy exponent.

Lemma 2 (Control of $g(\omega)$). *Let V be a Lévy measure, and define $\mathcal{A}_{\text{sym}} = \emptyset$ if V is symmetric and $\mathcal{A}_{\text{sym}} = \{1\}$ otherwise. For some $0 < p \leq q \leq 2$ let $\mathcal{A} = \{p, q\} \cup \mathcal{A}_{\text{sym}}$ and set $p_{\min} = \min \mathcal{A}$ and $p_{\max} = \max \mathcal{A}$. Then, if $V \in \mathcal{M}(p_{\min}, p_{\max})$, for the function*

$$g(\omega) = \int_{\mathbb{R} \setminus \{0\}} (e^{j\omega a} - 1 - j\omega a \mathbb{1}_{|a| < 1}) V(da), \quad (34)$$

there exist constants κ_1 and $\kappa_2 \geq 0$ such that, for all $(\omega_1, \omega_2) \in \mathbb{R}^2$,

$$|g(\omega_2) - g(\omega_1)| \leq \kappa_1 h_{p_{\min}}(\omega_1, \omega_2) + \kappa_2 h_{p_{\max}}(\omega_1, \omega_2). \quad (35)$$

Proof. We decompose $(g(\omega_1) - g(\omega_2))$ into 4 parts as

$$\begin{aligned} g(\omega_1) - g(\omega_2) &= \int_{|a| < 1} (\cos(a\omega_1) - \cos(a\omega_2)) V(da) \\ &\quad + \int_{|a| \geq 1} (\cos(a\omega_1) - \cos(a\omega_2)) V(da) \\ &\quad + j \int_{|a| < 1} (\sin(a\omega_1) - \sin(a\omega_2) - a(\omega_1 - \omega_2)) V(da) \\ &\quad + j \int_{|a| \geq 1} (\sin(a\omega_1) - \sin(a\omega_2)) V(da) \\ &= g_{\mathbb{R},0}(\omega_{1,2}) + g_{\mathbb{R},\infty}(\omega_{1,2}) + g_{\mathbb{S},0}(\omega_{1,2}) + g_{\mathbb{S},\infty}(\omega_{1,2}). \end{aligned}$$

To simplify the notations we introduce $\Delta = \frac{a}{2}(\omega_1 - \omega_2)$ and $\Sigma = \frac{a}{2}(\omega_1 + \omega_2)$. We can write that

$$\begin{cases} \cos(a\omega_1) - \cos(a\omega_2) &= -2 \sin(\Delta) \sin(\Sigma), \\ \sin(a\omega_1) - \sin(a\omega_2) &= 2 \sin(\Delta) \cos(\Sigma). \end{cases}$$

1. We start with $g_{\mathfrak{R},0}$ and use the fact that $|\sin x| \leq \min(1, |x|) \leq |x|^{\frac{p_{\max}}{2}}$, because $p_{\max} \leq 2$.

$$\begin{aligned}
|g_{\mathfrak{R},0}(\omega_{1,2})| &\leq 2 \int_{|a|<1} |\sin(\Delta) \sin(\Sigma)| V(da) \\
&\leq 2 \int_{|a|<1} |\Delta \Sigma|^{\frac{p_{\max}}{2}} V(da) \\
&= 2^{1-p_{\max}} \mu_{p_{\max}}^0(V) |\omega_1^2 - \omega_2^2|^{\frac{p_{\max}}{2}} \\
&\leq \max\left(2^{1-p_{\max}}, 2^{\frac{1-p_{\max}}{2}}\right) \mu_{p_{\max}}^0(V) h_{p_{\max}}(\omega_1, \omega_2),
\end{aligned}$$

where we used part (i) of Lemma 1 with $p = p_{\max}$ for the last inequality.

2. For $g_{\mathfrak{R},\infty}$, we use $|\sin x| \leq |x|^{\frac{p_{\min}}{2}}$ and part (i) of Lemma 1 with $p = p_{\min}$ to obtain

$$\begin{aligned}
|g_{\mathfrak{R},\infty}(\omega_{1,2})| &\leq 2 \int_{|a|\geq 1} |\sin(\Delta) \sin(\Sigma)| V(da) \\
&\leq 2 \int_{|a|\geq 1} |\Delta \Sigma|^{\frac{p_{\min}}{2}} V(da) \\
&= 2^{1-p_{\min}} \mu_{p_{\min}}^\infty(V) |\omega_1^2 - \omega_2^2|^{\frac{p_{\min}}{2}} \\
&\leq 2^{\frac{1-p_{\min}}{2}} \max\left(1, 2^{\frac{1-p_{\min}}{2}}\right) \mu_{p_{\min}}^\infty(V) h_{p_{\min}}(\omega_1, \omega_2).
\end{aligned}$$

3. If V is symmetric, then $g_{\mathfrak{S},0} = 0$ and we do not need any bounds. For asymmetric cases, we know that $p_{\max} = \max(q, 1)$. Here, we use the inequality $|x - \sin x| \leq 2|x|^{p_{\max}} \in [1, 2]$ for $1 \leq p_{\max} \leq 2$. Indeed, $|x - \sin(x)| = |x|(1 - \text{sinc}(x)) \leq |x| \min(2, |x|) \leq 2|x| \times |x|^{p_{\max}-1}$. By recalling $\cos x = (1 - 2\sin^2(\frac{x}{2}))$, we have that

$$\begin{aligned}
|g_{\mathfrak{S},0}(\omega_{1,2})| &= 2 \left| \int_{|a|<1} (\sin(\Delta) \cos(\Sigma) - \Delta) V(da) \right| \\
&= 2 \left| \int_{|a|<1} \left(\Delta - \sin(\Delta) + 2 \sin(\Delta) \sin^2\left(\frac{\Sigma}{2}\right) \right) V(da) \right| \\
&\leq 2 \int_{|a|<1} \left(2|\Delta|^{p_{\max}} + 2|\Delta|^{\frac{p_{\max}}{2}} |\Sigma/2|^{\frac{p_{\max}}{2}} \right) V(da) \\
&= 4\mu_{p_{\max}}^0(V) \left(|\omega_1 - \omega_2|^{p_{\max}} + |\omega_1^2 - \omega_2^2|^{\frac{p_{\max}}{2}} \right) \\
&\leq 2^{\frac{p_{\max}+5}{2}} \mu_{p_{\max}}^0(V) h_{p_{\max}}(\omega_1, \omega_2),
\end{aligned}$$

where we used part (i) of Lemma 1 for $p = p_{\max} \geq 1$.

4. Again, if V is symmetric, $g_{\mathfrak{S},\infty} = 0$. The construction of \mathcal{A} for asymmetric cases implies that $p_{\min} = \min(p, 1)$. By using the inequality $|\sin x| \leq |x|^{p_{\min}}$ and part (i) of Lemma 1, we get that

$$\begin{aligned}
|g_{\mathfrak{S},\infty}(\omega_{1,2})| &\leq 2 \int_{|a|\geq 1} |\sin(\Delta) \cos(\Sigma)| V(da) \\
&\leq 2 \int_{|a|\geq 1} |\Delta|^{p_{\min}} V(da) \\
&= 2\mu_{p_{\min}}^\infty(V) |\omega_1 - \omega_2|^{p_{\min}} \\
&\leq 2\mu_{p_{\min}}^\infty(V) h_{p_{\min}}(\omega_1, \omega_2).
\end{aligned}$$

We now just have to sum the four bounds to get the result. □

Corollary 1. *Under the same assumptions of Lemma 2,*

$$|g(\omega)| \leq \kappa_1 |\omega|^{p_{\min}} + \kappa_2 |\omega|^{p_{\max}}.$$

Proof. The result is obvious by setting $(\omega_1, \omega_2) = (\omega, 0)$ in Lemma 2. \square

We now focus on bounding the generalized Lévy exponent $G(\varphi) = \int_{\mathbb{R}^d} g(\varphi(\mathbf{r})) d\mathbf{r}$ with no Gaussian part.

Lemma 3 (Control of $G(\varphi)$). *Let V be a Lévy measure and define $\mathcal{A}_{\text{sym}} = \emptyset$ if V is symmetric and $\mathcal{A}_{\text{sym}} = \{1\}$ otherwise. For some $0 < p \leq q \leq 2$ let $\mathcal{A} = \{p, q\} \cup \mathcal{A}_{\text{sym}}$ and set $p_{\min} = \min \mathcal{A}$ and $p_{\max} = \max \mathcal{A}$. Then, if $V \in \mathcal{M}(p_{\min}, p_{\max})$, the functional*

$$G(\varphi) = \int_{\mathbb{R}^d} g(\varphi(\mathbf{r})) d\mathbf{r} \quad (36)$$

is well-defined on $L^{p_{\min}} \cap L^{p_{\max}}$ and there exist $\kappa_1, \kappa_2 \geq 0$ such that, for all φ and $\psi \in L^{p_{\min}} \cap L^{p_{\max}}$,

$$|G(\varphi) - G(\psi)| \leq \kappa_1 H_{p_{\min}}(\varphi, \psi) + \kappa_2 H_{p_{\max}}(\varphi, \psi). \quad (37)$$

Proof. we use Corollary 1 to prove that G is well-defined. Indeed,

$$\int_{\mathbb{R}^d} |g(\varphi(\mathbf{r}))| d\mathbf{r} \leq \kappa_1 \|\varphi\|_{p_{\min}}^{p_{\min}} + \kappa_2 \|\varphi\|_{p_{\max}}^{p_{\max}} < +\infty,$$

which proves that $g(\varphi) \in L^1$. Then, we apply Lemmas 1 and 2 to conclude that

$$\begin{aligned} |G(\varphi) - G(\psi)| &\leq \int_{\mathbb{R}^d} |g(\varphi(\mathbf{r})) - g(\psi(\mathbf{r}))| d\mathbf{r} \\ &\leq \kappa_1 \int_{\mathbb{R}^d} h_{p_{\max}}(\varphi(\mathbf{r}), \psi(\mathbf{r})) d\mathbf{r} + \kappa_2 \int_{\mathbb{R}^d} h_{p_{\max}}(\varphi(\mathbf{r}), \psi(\mathbf{r})) d\mathbf{r} \\ &\leq \kappa_1 H_{p_{\max}}(\varphi, \psi) + \kappa_2 H_{p_{\max}}(\varphi, \psi). \end{aligned}$$

\square

By including the Gaussian part, we now give the continuity condition in its general form for the characteristic functional of innovation processes.

Proposition 4 (Continuity of the characteristic functional). *Let f be a Lévy exponent with triplet (μ, σ^2, V) and let $0 < p \leq q \leq 2$. We define*

- $\mathcal{A}_{\text{sym}} = \emptyset$ if V is symmetric and $\{1\}$ otherwise,
- $\mathcal{A}_1 = \emptyset$ if $\mu = 0$ and $\{1\}$ otherwise,
- $\mathcal{A}_2 = \emptyset$ if $\sigma^2 = 0$ and $\{2\}$ otherwise,
- $\mathcal{A} = \{p, q\} \cup \mathcal{A}_{\text{sym}} \cup \mathcal{A}_1 \cup \mathcal{A}_2$,
- $p_{\min} = \min \mathcal{A}$ and $p_{\max} = \max \mathcal{A}$.

If $V \in \mathcal{M}(p_{\min}, p_{\max})$, then, the generalized Lévy exponent $F(\varphi) = \int_{\mathbb{R}^d} f(\varphi(\mathbf{r})) d\mathbf{r}$ is well-defined on $L^{p_{\min}} \cap L^{p_{\max}}$ and there exist $\nu_1, \nu_2 \geq 0$ such that, for all φ and $\psi \in L^{p_{\min}} \cap L^{p_{\max}}$, we have that

$$|F(\varphi) - F(\psi)| \leq \nu_1 H_{p_{\min}}(\varphi, \psi) + \nu_2 H_{p_{\max}}(\varphi, \psi). \quad (38)$$

This implies that $F(\varphi)$ is continuous over $L^{p_{\min}} \cap L^{p_{\max}}$.

Proof. We use Lemma 3 to justify that $G(\varphi) = \int_{\mathbb{R}^d} g(\varphi(\mathbf{r})) d\mathbf{r}$ is well-defined over $L^{p_{\min}} \cap L^{p_{\max}}$ and there exist $\kappa_1, \kappa_2 > 0$ such that

$$|G(\varphi) - G(\psi)| \leq \kappa_1 H_{p_{\min}}(\varphi, \psi) + \kappa_2 H_{p_{\max}}(\varphi, \psi),$$

where $\varphi, \psi \in L^{p_{\min}} \cap L^{p_{\max}}$. Also, the inclusion of \mathcal{A}_1 and \mathcal{A}_2 in the definition of \mathcal{A} (and therefore p_{\min}, p_{\max}), imposes bounds on the Gaussian part as

$$\begin{aligned} |F(\varphi) - F(\psi)| &= \left| j\mu \left(\int (\varphi - \psi) \right) - \frac{\sigma^2}{2} \|\varphi - \psi\|_2^2 + G(\varphi) - G(\psi) \right| \\ &\leq |\mu| H_1(\varphi, \psi) + \frac{\sigma^2}{2} H_2(\varphi, \psi) + \kappa_1 H_{p_{\min}}(\varphi, \psi) + \kappa_2 H_{p_{\max}}(\varphi, \psi) \\ &\leq \nu_1 H_{p_{\min}}(\varphi, \psi) + \nu_2 H_{p_{\max}}(\varphi, \psi). \end{aligned}$$

To validate the last inequality, note that, if $\mu \neq 0$ ($\sigma \neq 0$), then $1 \in [p_{\min}, p_{\max}]$ ($2 \in [p_{\min}, p_{\max}]$). Hence, using Lemma 1, $H_1(\varphi, \psi)$ ($H_2(\varphi, \psi)$) can be upper-bounded by a linear combination of $H_{p_{\min}}(\varphi, \psi)$ and $H_{p_{\max}}(\varphi, \psi)$.

Finally, to ensure the continuity of F , we point out the fact that $H_p(\varphi_n, \varphi) \rightarrow 0$ if $\varphi_n \rightarrow \varphi$ in L^p . \square

4 Applications for Particular Classes of Operators

We use the previous results, mainly Theorem 5, for particular classes of differential operators L (classical or fractional). For each case, we summarize the hypotheses on w required in our results to define the sparse process s with $Ls = w$.

We first review the necessary steps for demonstrating the existence of s . As mentioned in Section 3.2, the interpretation of the innovation model $Ls = w$ is based on a characteristic functional of the form

$$\widehat{\mathcal{P}}_s(\varphi) = \exp \left(\int_{\mathbb{R}^d} f(L^{*-1}\varphi(\mathbf{r}))d\mathbf{r} \right). \quad (39)$$

Let L be a linear operator defined on \mathcal{S} (or a larger space) that has the adjoint operator L^* such that its adjoint admits a linear left inverse $L^{*-1}: \mathcal{S} \mapsto L^p \cap L^q$. Then, the characteristic functional of Ls is $\widehat{\mathcal{P}}_{Ls}(\varphi) = \exp \left(\int_{\mathbb{R}^d} f(L^{*-1}L^*\varphi(\mathbf{r}))d\mathbf{r} \right) = \exp \left(\int_{\mathbb{R}^d} f(\varphi(\mathbf{r}))d\mathbf{r} \right)$. In other words, the operator L whitens the generalized process s .

4.1 Self-Similar Sparse Processes

We are interested in defining generalized stochastic processes s such that $Ls = w$, where $L = (-\Delta)^{\gamma/2}$ is the fractional Laplacian operator of order $\gamma > 0$. Such processes are called second-order self-similar because their correlation structures are invariant to similarity transformations, due to the homogeneity of the fractional Laplacian operator. In the Gaussian case, self-similarity is intimately tied to fractional Brownian motions (fBm) [MN68]. The link between innovation models arising from fractional Laplacian operators and fBm is studied in [TVU09]. This indicates implicitly that such processes are special cases of the present framework. Here, by applying the fundamental results in [SU12], we extend the definition of Gaussian self-similar processes to the larger class of self-similar processes with infinite-divisible distributions derived from Lévy noises.

The fractional Laplacian operator $(-\Delta)^{\gamma/2}$ is defined for $\varphi \in \mathcal{S}$ by $(-\Delta)^{\gamma/2}\varphi = \mathcal{F}^{-1}(\|\cdot\|^\gamma \mathcal{F}\varphi)$ where \mathcal{F} (\mathcal{F}^{-1} , respectively) is the Fourier transform (the inverse Fourier transform, respectively). It is linear, continuous from \mathcal{S} to C^∞ , rotational-, shift-, scale-invariant, and self-adjoint. Thus, we need to find its linear left inverse operator(s). For $0 < \gamma < d$, its natural (left and right) inverse is the well-known Riesz potential I_γ . An extension for $\gamma > d$, $\gamma \notin \mathbb{N}$, called generalized Riesz potential, is introduced in [SU12]. The main results concerning such operators can be found in Theorem 1.1 of [SU12] which is summarized in Theorem 6.

Theorem 6 (Shift-invariant left inverse of $(-\Delta)^{\gamma/2}$). *Let $\gamma > 0$ with $(\gamma - d) \notin \mathbb{N}$. The operator I_γ with frequency response $\|\omega\|^{-\gamma}$ is the unique linear and continuous operator from \mathcal{S} to \mathcal{S}' that is shift- and scale-invariant. It is also a left inverse of the fractional Laplacian operator on \mathcal{S} , which implies that*

$$\forall \varphi \in \mathcal{S}, \quad I_\gamma(-\Delta)^{\gamma/2}\varphi = \varphi. \quad (40)$$

In general, the output range of operators I_γ on $\varphi \in \mathcal{S}$ is not restricted to \mathcal{S} or even to an L^p space. However, we can confine the range by limiting γ . More precisely, by considering a generalization of the Hardy-Littlewood-Sobolev inequality for $1 \leq p \leq +\infty$, we can show that

$$I_\gamma(\mathcal{S}) \subset L^p \Leftrightarrow 0 < \gamma < d(1 - 1/p). \quad (41)$$

Consequently,

- For $\gamma < d$, we have that $I_\gamma(\mathcal{S}) \subset L^p$ iff. $p > \frac{d}{d-\gamma} > 1$.
- For $d < \gamma \notin \mathbb{N}$, we have that $I_\gamma(\mathcal{S}) \not\subset L^p$ for all $p \geq 1$.

Suppose now that $\gamma \notin \mathbb{N}$. Following [SU12], we are able to define a correction of I_γ from \mathcal{S} to L^p for all fixed $p \geq 1$ such that $(\gamma - d(1 - 1/p)) \notin \mathbb{N}$. We denote the set of forbidden values of p by $A(d, \gamma)$, which is given by

$$A(d, \gamma) = \begin{cases} \left\{ \frac{d}{d+1-\epsilon(\gamma)-k}, k \in \{0, \dots, \lfloor \gamma \rfloor\} \right\} & \text{if } \gamma < d, \\ \left\{ \frac{d}{d+1-\epsilon(\gamma)-k}, k \in \{0, \dots, d\} \right\} & \text{if } \gamma > d. \end{cases} \quad (42)$$

The set $A(d, \gamma)$ contains $k(d, \gamma) = \min(\lfloor \gamma \rfloor + 1, d)$ elements. Therefore, $[1, +\infty] \setminus A(d, \gamma)$ is composed of $k(d, \gamma) + 1$ intervals given by

$$C(d, \gamma, k) = \begin{cases} \left[1, \frac{d}{d-\epsilon(\gamma)} \right) & \text{if } k = 0, \\ \left(\frac{d}{d-k+1-\epsilon(\gamma)}, \frac{d}{d-k-\epsilon(\gamma)} \right) & \text{if } k \in \{1, \dots, k(d, \gamma) - 1\}, \\ \left(\frac{d}{d-k+1-\epsilon(\gamma)}, +\infty \right] & \text{if } k = k(d, \gamma). \end{cases} \quad (43)$$

For instance, if $d = 1$, then, $k(1, \gamma) = 1$ and there is only one forbidden value, with $p = \frac{1}{1-\epsilon(\gamma)}$. Also, the two intervals are $C(1, \gamma, 0) = [1, \frac{1}{1-\epsilon(\gamma)})$ and $C(1, \gamma, 1) = (\frac{1}{1-\epsilon(\gamma)}, +\infty]$. Similarly, for $d = 2$, there are two intervals if $\gamma < 1$ and three otherwise.

We modify I_γ to guarantee some L^p stability, for $k \in \{0, \dots, k(d, \gamma)\}$, by

$$\mathcal{F}(I_{\gamma,k}\{\varphi\})(\omega) = \|\omega\|^{-\gamma} \left(\mathcal{F}\varphi(\omega) - \sum_{|\mathbf{j}| \leq \lfloor \gamma \rfloor - k} \frac{\partial^{\mathbf{j}} \mathcal{F}\varphi(\mathbf{0})}{\mathbf{j}!} \omega^{\mathbf{j}} \right). \quad (44)$$

Note that such operators are no longer shift-invariant. This can be cast as the cost of obtaining an L^p -stable operator.

Proposition 5 (L^p -stable left inverse of $(-\Delta)^{\gamma/2}$). *Let $d \in \mathbb{N}^*$, $\gamma \in (0, +\infty) \setminus (\mathbb{N} + d)$, and $k \in \{0, \dots, k(d, \gamma)\}$.*

- *The operator $I_{\gamma,k}$ is continuous linear and scale-invariant from \mathcal{S} to L^p , for all $p \in C(d, \gamma, k)$. Moreover, it is a left inverse of the Laplacian operator $(-\Delta)^{\gamma/2}$.*
- *For fixed $p \in C(d, \gamma, k)$, $I_{\gamma,k}$ is the unique linear and scale-invariant left inverse of $(-\Delta)^{\gamma/2}$ from \mathcal{S} to L^p .*
- *If p and q are in distinct $C(d, \gamma, k)$ sets, then the Laplacian operator $(-\Delta)^{\gamma/2}$ has no linear and scale-invariant left inverse from \mathcal{S} to $L^p \cap L^q$.*

Proof. The first two claims are direct rewritings of Theorem 1.2 in [SU12] by noting that, for $p \in [1, +\infty] \setminus A(\gamma)$,

$$p \in C(d, \gamma, k) \Leftrightarrow \left\lfloor \gamma - d \left(1 - \frac{1}{p} \right) \right\rfloor = \lfloor \gamma \rfloor - k. \quad (45)$$

The last claim follows from the uniqueness property and states that the conditions for restricting the range to L^p and L^q are incompatible. \square

We are now able to give admissibility conditions between γ and a Lévy exponent f to define processes whitened by fractional Laplacian.

Proposition 6. *Let $\gamma \in (0, +\infty) \setminus \mathbb{N}$ and f be a Lévy exponent with triplet (μ, σ^2, V) . Define p_{\min} and p_{\max} as in Theorem 5 and let k be such that p_{\min} and $p_{\max} \in C(d, \gamma, k)$. Then, there exists a generalized stochastic process s with*

$$\widehat{\mathcal{P}}_s(\varphi) = \exp \left(\int_{\mathbb{R}^d} f(I_{\gamma,k}\varphi(\mathbf{r})) d\mathbf{r} \right) \quad (46)$$

for $\varphi \in \mathcal{S}$. *The process s is a broadly self-similar sparse process (self-similar Gaussian process, respectively) if $V \neq 0$ (if $V \equiv 0$, respectively).*

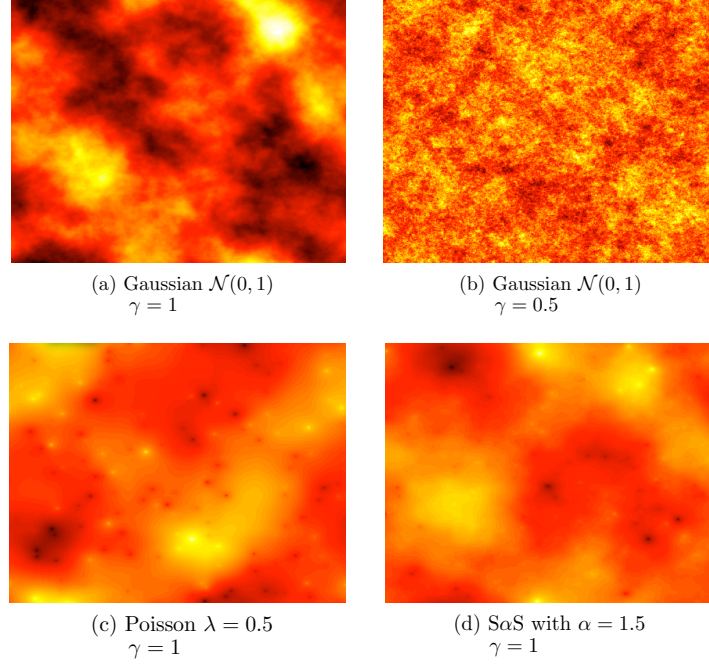


Figure 2: Self-similar processes

Proposition 6 can be interpreted as follows: there exists a process s , such that $\langle s, \varphi \rangle = \langle w, I_{\gamma,k} \varphi \rangle = \langle I_{\gamma,k}^* w, \varphi \rangle$. In other words, $(-\Delta)^{\gamma/2} s = w$.

Proof. We apply Theorem 5 with $T = I_{\gamma,k}$, which is continuous on $L^{p_{\min}} \cap L^{p_{\max}}$, according to Proposition 5. \square

We examine the construction of the innovation models in Proposition 6 for Lévy-Schwartz measures V with finite first-order moment $\mu_1^\infty(V) < +\infty$.

- **Gaussian case.** Suppose that $(\gamma - \frac{d}{2}) \notin \mathbb{N}$. Then, there exists $k = \lfloor d/2 + 1 - \epsilon(\gamma) \rfloor$ such that $I_{\gamma,k}$ is continuous from \mathcal{S} to L^2 . Thus, according to Proposition 6, the functional $\widehat{\mathcal{P}}_s(\varphi) = \exp\left(-\frac{\sigma^2 \|I_{\gamma,k} \varphi\|_2^2}{2}\right)$ defines a process s which is whitened by $(-\Delta)^{\gamma/2}$.
- **Laplace case.** The Lévy measure V_L of a Laplacian law verifies $V_L \in \mathcal{M}(1,1)$. Let $p_{\min} = p_{\max} = 1$. Proposition 6 applies with the operator $I_{\gamma,0}$ for $\gamma \notin \mathbb{N}$.
- **Compound Poisson case.** Suppose that $V \in \mathcal{M}(1,0)$. Then, as in the Laplace case, the operator $I_{\gamma,0}$ is admissible for all $\gamma \notin \mathbb{N}$.
- **SaS.** Let $1 \leq \alpha < 2$ and $\gamma \notin \mathbb{N}$ with $(\gamma - d(1 - 1/\alpha)) \notin \mathbb{N}$. Then, there exists $k = \lfloor d(1 - 1/\alpha) + 1 - \epsilon(\gamma) \rfloor$ such that $I_{\gamma,k}(\mathcal{S}) \subset L^\alpha$. According to Proposition 6, there exists s with $\widehat{\mathcal{P}}_s(\varphi) = \exp(-\|I_{\gamma,k} \varphi\|_\alpha^\alpha)$ and $(-\Delta)^{\gamma/2} s = w$ with w an α -stable innovation process.

We depict in Figure 2 some examples of self-similar processes in dimension $d = 2$. Dark intensities correspond to the highest values of the simulated process, while bright ones correspond to the smallest.

4.2 Directional Sparse Processes

Our goal is to define directional stochastic processes on \mathcal{S}' using oriented differential operators. This consists of defining proper left inverse operators for derivative operators of the form $D_{\mathbf{u}} \varphi = \langle \nabla \varphi, \mathbf{u} \rangle = u_1 D_1 \varphi + \dots + u_d D_d \varphi$, where \mathbf{u} stands for the direction. For this purpose, we extend the one-dimensional results of [UTS11] to higher dimensions. We start with first-order operators $L = D_{\mathbf{u}} - \alpha \text{Id}$ with $\alpha \in \mathbb{C}$ and $\mathbf{u} \in \mathbb{R}^d$.

We denote by (\mathbf{e}_k) the canonical basis of \mathbb{R}^d . For $\mathbf{u} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, $p_{\mathbf{u}^\perp}(\mathbf{r}) = \mathbf{r} - \frac{\langle \mathbf{u}, \mathbf{r} \rangle}{\|\mathbf{u}\|_2^2} \mathbf{u}$ is the orthogonal projection on $\mathbf{u}^\perp = \{\mathbf{v} \mid \langle \mathbf{u}, \mathbf{v} \rangle = 0\}$. Recall that $\|\mathbf{r}\|_2^2 = \langle \mathbf{u}, \mathbf{r} \rangle^2 \|\mathbf{u}\|_2^2 + \|p_{\mathbf{u}^\perp}(\mathbf{r})\|_2^2$. Since $\mathbf{D}_{\mathbf{u}} = \|\mathbf{u}\|_2 \mathbf{D}_{\mathbf{u}/\|\mathbf{u}\|_2}$, we assume now that $\|\mathbf{u}\|_2 = 1$, without loss of generality.

4.2.1 Left Inverse Operators of $\mathbf{L} = \mathbf{D}_{\mathbf{u}} - \alpha \text{Id}$

Let $\mathbf{u} \in \mathbb{R}^d$ with norm $\|\mathbf{u}\|_2 = 1$. We separately investigate the cases of $\Re(\alpha) \neq 0$ and $\Re(\alpha) = 0$, as they result in stable and marginally stable left inverses, respectively.

Left inverse operators in the stable case

Since the case of $\Re(\alpha) > 0$ is very similar to $\Re(\alpha) < 0$, we first study the causal case. Therefore, we assume that $\Re(\alpha) < 0$ and we define $\rho_{\mathbf{u}, \alpha}$ by

$$\langle \rho_{\mathbf{u}, \alpha}, \varphi \rangle = \int_0^{+\infty} e^{\alpha t} \varphi(t\mathbf{u}) dt. \quad (47)$$

We further define the operator $I_{\mathbf{u}, \alpha}$ on \mathcal{S} as $I_{\mathbf{u}, \alpha} = \rho_{\mathbf{u}, \alpha} * \varphi$. In one dimension, $\rho_{\mathbf{u}, \alpha}$ is a causal exponential function in the classical sense as was introduced in [UTS11]. However, for $d \geq 2$, it is a generalized function.

Proposition 7. *The continuous operator $I_{\mathbf{u}, \alpha}$ is LSI and continuous from \mathcal{S} to \mathcal{S} . Furthermore, it is the inverse of the partial differential operator $(\mathbf{D}_{\mathbf{u}} - \alpha \text{Id})$ on \mathcal{S} , meaning that*

$$I_{\mathbf{u}, \alpha}(\mathbf{D}_{\mathbf{u}} - \alpha \text{Id})\varphi = (\mathbf{D}_{\mathbf{u}} - \alpha \text{Id})I_{\mathbf{u}, \alpha}\varphi = \varphi, \quad \forall \varphi \in \mathcal{S}. \quad (48)$$

Proof. First, because $|\langle \rho_{\mathbf{u}, \alpha}, \varphi \rangle| \leq \|\varphi\|_\infty \int_0^{+\infty} |e^{\alpha t}| dt = \|\varphi\|_\infty / (-\Re(\alpha))$, we have that $\rho_{\mathbf{u}, \alpha} \in \mathcal{S}'$. This confirms that $I_{\mathbf{u}, \alpha}$ is well-defined on \mathcal{S} .

The derivative of $\rho_{\mathbf{u}, \alpha}$ in the sense of generalized functions is

$$\begin{aligned} \langle \mathbf{D}_{\mathbf{u}} \rho_{\mathbf{u}, \alpha}, \varphi \rangle &= -\langle \rho_{\mathbf{u}, \alpha}, \mathbf{D}_{\mathbf{u}} \varphi \rangle \\ &= -\int_0^{+\infty} e^{\alpha t} \{\mathbf{D}_{\mathbf{u}} \varphi\}(t\mathbf{u}) dt \\ &= -[\varphi(t\mathbf{u}) e^{\alpha t}]_0^{+\infty} + \alpha \langle \rho_{\mathbf{u}, \alpha}, \varphi \rangle \\ &\quad \text{[using an integration by parts]} \\ &= \varphi(0) + \alpha \langle \rho_{\mathbf{u}, \alpha}, \varphi \rangle, \end{aligned}$$

meaning that $(\mathbf{D}_{\mathbf{u}} - \alpha \text{Id})\rho_{\mathbf{u}, \alpha} = \delta$. Consequently, in Fourier domain, we have that

$$\begin{aligned} 1 &= \widehat{\mathbf{D}_{\mathbf{u}} \rho_{\mathbf{u}, \alpha}}(\boldsymbol{\omega}) - \alpha \widehat{\rho_{\mathbf{u}, \alpha}}(\boldsymbol{\omega}) \\ &= (\mathbf{j} \langle \boldsymbol{\omega}, \mathbf{u} \rangle - \alpha) \widehat{\rho_{\mathbf{u}, \alpha}}(\boldsymbol{\omega}). \end{aligned}$$

This implies

$$\widehat{\rho_{\mathbf{u}, \alpha} * \varphi}(\boldsymbol{\omega}) = \widehat{\rho_{\mathbf{u}, \alpha}}(\boldsymbol{\omega}) \widehat{\varphi}(\boldsymbol{\omega}) = \frac{\widehat{\varphi}(\boldsymbol{\omega})}{\mathbf{j} \langle \boldsymbol{\omega}, \mathbf{u} \rangle - \alpha}.$$

Consequently, $\widehat{\rho_{\mathbf{u}, \alpha} * \varphi}$ and $\rho_{\mathbf{u}, \alpha} * \varphi$ belong to \mathcal{S} . Moreover, we know that the LSI operator $\varphi \mapsto u * \varphi$ for $u \in \mathcal{S}'$, is continuous from \mathcal{S} into itself iff. $u \in \mathcal{O}'_C$ or, equivalently, iff. $\widehat{u} \in \mathcal{O}_M$, the space of slowly increasing and infinitely differentiable functions (see [Sch66] for more details). Since in our case $\widehat{\rho_{\mathbf{u}, \alpha}} \in \mathcal{O}_M$, we conclude that $I_{\mathbf{u}, \alpha}$ is continuous.

Finally, we can write that

$$\begin{aligned} (\mathbf{D}_{\mathbf{u}} - \alpha \text{Id})I_{\mathbf{u}, \alpha}\varphi &= (\mathbf{D}_{\mathbf{u}} - \alpha \text{Id})(\rho_{\mathbf{u}, \alpha} * \varphi) = ((\mathbf{D}_{\mathbf{u}} - \alpha \text{Id})\rho_{\mathbf{u}, \alpha}) * \varphi = \delta * \varphi = \varphi, \\ I_{\mathbf{u}, \alpha}(\mathbf{D}_{\mathbf{u}} - \alpha \text{Id})\varphi &= \rho_{\mathbf{u}, \alpha} * ((\mathbf{D}_{\mathbf{u}} - \alpha \text{Id})\varphi) = ((\mathbf{D}_{\mathbf{u}} - \alpha \text{Id})\rho_{\mathbf{u}, \alpha}) * \varphi = \varphi. \end{aligned}$$

□

Following [UTS11], we can transpose this result for $\Re(\alpha) > 0$ (anti-causal case) by defining $\rho_{\mathbf{u}, -\alpha}(\mathbf{r}) = \rho_{\mathbf{u}, -\alpha}(-\mathbf{r})$. With this choice, we can show in a similar way that Proposition 7 also holds for $\Re(\alpha) > 0$.

Left inverse operators in the marginally stable case

Suppose now that $\alpha = j\omega_0$ is purely imaginary, with $\Re(\alpha) = 0$. The natural candidate for $(D_{\mathbf{u}} - j\omega_0 \text{Id})^{-1}$ is again the convolution operator defined by the kernel $\rho_{\mathbf{u}, j\omega_0}$ where

$$\langle \rho_{\mathbf{u}, j\omega_0}, \varphi \rangle = \int_0^{+\infty} e^{j\omega_0 t} \varphi(t\mathbf{u}) dt.$$

In other words,

$$I_{\mathbf{u}, j\omega_0} \varphi(\mathbf{r}) = e^{j\omega_0 \langle \mathbf{r}, \mathbf{u} \rangle} \int_{-\infty}^{\langle \mathbf{r}, \mathbf{u} \rangle} e^{-j\omega_0 \tau} \varphi(p_{\mathbf{u}^\perp}(\mathbf{r}) + \tau \mathbf{u}) d\tau. \quad (49)$$

The adjoint of $I_{\mathbf{u}, j\omega_0}$ is given by

$$I_{\mathbf{u}, j\omega_0}^* \varphi(\mathbf{r}) = (\rho_{\mathbf{u}, j\omega_0}(-\cdot) * \varphi)(\mathbf{r}) = e^{-j\omega_0 \langle \mathbf{r}, \mathbf{u} \rangle} \int_{\langle \mathbf{r}, \mathbf{u} \rangle}^{+\infty} e^{j\omega_0 \tau} \varphi(p_{\mathbf{u}^\perp}(\mathbf{r}) + \tau \mathbf{u}) d\tau. \quad (50)$$

These $I_{\mathbf{u}, j\omega}$ operators are shift-invariant. However, their impulse responses are not rapidly decreasing (their Fourier transforms are not in \mathcal{O}_M). Consequently, they are not stable and cannot define valid characteristic functionals in (39). Here, we propose a modification inspired by [UTS11] to overcome the instability problem. We define

$$\begin{aligned} J_{\mathbf{u}, \omega_0} \varphi(\mathbf{r}) &= I_{\mathbf{u}, j\omega_0} \varphi(\mathbf{r}) - \{I_{\mathbf{u}, j\omega_0} \varphi\}(p_{\mathbf{u}^\perp}(\mathbf{r})) e^{j\omega_0 \langle \mathbf{r}, \mathbf{u} \rangle} \\ &= e^{j\omega_0 \langle \mathbf{r}, \mathbf{u} \rangle} \int_0^{\langle \mathbf{r}, \mathbf{u} \rangle} e^{-j\omega_0 \tau} \varphi(p_{\mathbf{u}^\perp}(\mathbf{r}) + \tau \mathbf{u}) d\tau. \end{aligned} \quad (51)$$

The modified operator $J_{\mathbf{u}, \omega_0}$ is continuous from \mathcal{S} to C^∞ and is a right-inverse of $(D_{\mathbf{u}} - j\omega_0 \text{Id})$. Indeed, $(I_{\mathbf{u}, j\omega_0} \varphi - J_{\mathbf{u}, \omega_0} \varphi)(\mathbf{r}) = I_{\mathbf{u}, j\omega_0} \varphi(p_{\mathbf{u}^\perp}(\mathbf{r})) e^{j\omega_0 \langle \mathbf{r}, \mathbf{u} \rangle} \in \text{Ker}(D_{\mathbf{u}} - j\omega_0 \text{Id})$. We claim that its adjoint is given by

$$J_{\mathbf{u}, \omega_0}^* \varphi(\mathbf{r}) = I_{\mathbf{u}, j\omega_0}^* \varphi(\mathbf{r}) - \mathbb{1}_{\langle \mathbf{r}, \mathbf{u} \rangle \leq 0} e^{-j\omega_0 \langle \mathbf{r}, \mathbf{u} \rangle} \int_{-\infty}^{+\infty} e^{j\omega_0 \tau} \varphi(p_{\mathbf{u}^\perp}(\mathbf{r}) + \tau \mathbf{u}) d\tau. \quad (53)$$

We have now to show that

$$\forall \varphi, \psi \in \mathcal{S}, \quad \langle J_{\mathbf{u}, \omega_0} \varphi, \psi \rangle = \langle \varphi, J_{\mathbf{u}, \omega_0}^* \psi \rangle. \quad (54)$$

We denote

$$A(\varphi, \psi) = \langle I_{\mathbf{u}, j\omega_0} \varphi(p_{\mathbf{u}^\perp}(\mathbf{r})) e^{j\omega_0 \langle \mathbf{r}, \mathbf{u} \rangle}, \psi(\mathbf{r}) \rangle \quad (55)$$

$$B(\varphi, \psi) = \langle \varphi(\mathbf{r}), \mathbb{1}_{\langle \mathbf{r}, \mathbf{u} \rangle \leq 0} e^{-j\omega_0 \langle \mathbf{r}, \mathbf{u} \rangle} \int_{-\infty}^{+\infty} e^{j\omega_0 \tau} \psi(p_{\mathbf{u}^\perp}(\mathbf{r}) + \tau \mathbf{u}) d\tau \rangle. \quad (56)$$

The Equation (54) is equivalent (as we can see from (51) and (53)) to $A(\varphi, \psi) = B(\varphi, \psi)$. To show this, we denote $(\mathbf{u}_1 = \mathbf{u}, \mathbf{u}_2, \dots, \mathbf{u}_d)$ an orthonormal basis of \mathbb{R}^d . Especially, we have $p_{\mathbf{u}^\perp}(\mathbf{r}) = \langle \mathbf{r}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \dots + \langle \mathbf{r}, \mathbf{u}_d \rangle \mathbf{u}_d$. Then,

$$\begin{aligned} A(\varphi, \psi) &= \int_{\mathbb{R}^d} \left(e^{j\omega_0 \langle \mathbf{r}, \mathbf{u} \rangle} \int_{\mathbb{R}} \mathbb{1}_{\tau \leq 0} e^{-j\omega_0 \tau} \psi(p_{\mathbf{u}^\perp}(\mathbf{r}) + \tau \mathbf{u}) d\tau \right) \varphi(\mathbf{r}) d\mathbf{r} \\ &= \int_{\mathbb{R}^{d+1}} \mathbb{1}_{\tau \leq 0} e^{j\omega_0 r_1} e^{-j\omega_0 \tau} \varphi(\tau \mathbf{u}_1 + r_2 \mathbf{u}_2 + \dots + r_d \mathbf{u}_d) \psi(r_1 \mathbf{u}_1 + r_2 \mathbf{u}_2 + \dots + r_d \mathbf{u}_d) d\tau dr_1 \dots dr_d \\ &\quad [\text{because } (\mathbf{u}_1, \dots, \mathbf{u}_d) \text{ is an orthonormal basis}] \\ &= \int_{\mathbb{R}^{d+1}} \mathbb{1}_{r'_1 \leq 0} e^{j\omega_0 \tau'} e^{-j\omega_0 r'_1} \varphi(r'_1 \mathbf{u}_1 + r_2 \mathbf{u}_2 + \dots + r_d \mathbf{u}_d) \psi(\tau' \mathbf{u}_1 + r_2 \mathbf{u}_2 + \dots + r_d \mathbf{u}_d) d\tau' dr'_1 \dots dr_d \\ &\quad [\text{using the change of variables } (r_1, \tau) = (\tau', r'_1)] \\ &= \int_{\mathbb{R}^d} \varphi(\mathbf{r}) \left(\mathbb{1}_{\langle \mathbf{r}, \mathbf{u} \rangle \leq 0} e^{-j\omega_0 \langle \mathbf{r}, \mathbf{u} \rangle} \int_{\mathbb{R}} e^{j\omega_0 \tau'} \psi(\tau' \mathbf{u} + p_{\mathbf{u}^\perp}(\mathbf{r})) d\tau' \right) d\mathbf{r} \\ &= B(\varphi, \psi). \end{aligned}$$

Denote the modulation operators by $M_{\mathbf{u},\omega_0}\varphi(\mathbf{r}) = e^{i\omega_0\langle\mathbf{r},\mathbf{u}\rangle}\varphi(\mathbf{r})$. Then it is not difficult to check that

$$\begin{aligned} I_{\mathbf{u},j\omega_0}\varphi &= M_{\mathbf{u},\omega_0}I_{\mathbf{u},0}M_{\mathbf{u},-\omega_0}\varphi, \\ I_{\mathbf{u},j\omega_0}^*\varphi &= M_{\mathbf{u},-\omega_0}I_{\mathbf{u},0}^*M_{\mathbf{u},\omega_0}\varphi, \\ J_{\mathbf{u},\omega_0}\varphi &= M_{\mathbf{u},\omega_0}J_{\mathbf{u},0}M_{\mathbf{u},-\omega_0}\varphi, \\ J_{\mathbf{u},\omega_0}^*\varphi &= M_{\mathbf{u},-\omega_0}J_{\mathbf{u},0}^*M_{\mathbf{u},\omega_0}\varphi. \end{aligned} \quad (57)$$

Note that $J_{\mathbf{u},\omega_0}$ preserves the regularity, with $J_{\mathbf{u},\omega_0}\varphi \in C^\infty$. On the contrary, its adjoint creates discontinuities along the hyperplane $\langle\mathbf{r},\mathbf{u}\rangle = 0$, while it preserves the decay properties, as we can see in Proposition 8.

Proposition 8 (Properties of $J_{\mathbf{u},\omega_0}^*$). *The following properties hold for the adjoint operator $J_{\mathbf{u},\omega_0}^*$ defined in (53):*

- The adjoint operator $J_{\mathbf{u},\omega_0}^*$ is continuous from $L^{\infty,\alpha}$ to $L^{\infty,\alpha-1}$ for $\alpha > 1$.
- The adjoint operator $J_{\mathbf{u},\omega_0}^*$ is linear and continuous from \mathcal{R} into itself, and it is a left inverse of the operator $(D_{\mathbf{u}} - j\omega_0\text{Id})^*$ on \mathcal{S} .

Proof. Because of the modulation equalities in (57), we only need to prove the claims for $\omega_0 = 0$. Let $\mathbf{r} \in \{\langle\mathbf{r},\mathbf{u}\rangle \geq 0\}$ and $\alpha > 1$. Then,

$$\begin{aligned} |J_{\mathbf{u},0}^*\varphi(\mathbf{r})| &= \left| \int_{\langle\mathbf{r},\mathbf{u}\rangle}^{+\infty} \varphi(p_{\mathbf{u}^\perp}(\mathbf{r}) + \tau\mathbf{u}) d\tau \right| \\ &\leq \|\varphi\|_{\infty,\alpha} \int_{\langle\mathbf{r},\mathbf{u}\rangle}^{+\infty} \frac{d\tau}{1 + \|p_{\mathbf{u}^\perp}(\mathbf{r}) + \tau\mathbf{u}\|_2^\alpha} \\ &= \|\varphi\|_{\infty,\alpha} \int_{\langle\mathbf{r},\mathbf{u}\rangle}^{+\infty} \frac{d\tau}{1 + \sqrt{\|p_{\mathbf{u}^\perp}(\mathbf{r})\|_2^2 + \tau^2}^\alpha} \\ &\quad [\text{using orthogonality of } p_{\mathbf{u}^\perp}(\mathbf{r}) \text{ and } \tau\mathbf{u}] \\ &\leq \|\varphi\|_{\infty,\alpha} \int_{\langle\mathbf{r},\mathbf{u}\rangle}^{+\infty} \frac{d\tau}{1 + (2^{-1/2}(\|p_{\mathbf{u}^\perp}(\mathbf{r})\|_2 + \tau))^\alpha} \\ &\quad [\text{using concavity of } \sqrt{\cdot}] \\ &\leq \|\varphi\|_{\infty,\alpha} \int_{(\langle\mathbf{r},\mathbf{u}\rangle + \|p_{\mathbf{u}^\perp}(\mathbf{r})\|_2)/\sqrt{2}}^{+\infty} \frac{d\nu}{1 + \nu^\alpha} \\ &\leq \frac{2^{3/2}\alpha}{\alpha-1} \|\varphi\|_{\infty,\alpha} \frac{1}{1 + 2^{-(\alpha-1)/2}(\langle\mathbf{r},\mathbf{u}\rangle + \|p_{\mathbf{u}^\perp}(\mathbf{r})\|_2)^{\alpha-1}} \\ &\quad [\text{using } \int_x^{+\infty} \frac{d\nu}{1 + \nu^\alpha} \leq \frac{2\alpha}{(\alpha-1)(1+x^{\alpha-1})} \text{ for } x \geq 0] \\ &\leq \frac{2^{3/2}\alpha}{\alpha-1} \|\varphi\|_{\infty,\alpha} \frac{1}{1 + 2^{-(\alpha-1)/2}\|\mathbf{r}\|_2^{\alpha-1}} \\ &\quad [\text{using } \langle\mathbf{r},\mathbf{u}\rangle + \|p_{\mathbf{u}^\perp}(\mathbf{r})\|_2 \geq \sqrt{\langle\mathbf{r},\mathbf{u}\rangle^2 + \|p_{\mathbf{u}^\perp}(\mathbf{r})\|_2^2} = \|\mathbf{r}\|_2]. \end{aligned}$$

Finally, we remark that $(1 + \|\mathbf{r}\|_2^{\alpha-1}) \leq 2^{(\alpha-1)/2}(1 + 2^{-(\alpha-1)/2}\|\mathbf{r}\|_2^{\alpha-1})$, which yields

$$|J_{\mathbf{u},0}^*\varphi(\mathbf{r})(1 + \|\mathbf{r}\|_2^{\alpha-1})| \leq C_\alpha \|\varphi\|_{\infty,\alpha}.$$

The same inequality holds for $\langle\mathbf{r},\mathbf{u}\rangle < 0$ which ensures the continuity from $L^{\infty,\alpha}$ to $L^{\infty,\alpha-1}$ for $\alpha > 1$. Because $\mathcal{R} = \bigcap_{\alpha>0} L^{\infty,\alpha}$, the previous bounds for all $\alpha > 1$ imply that $J_{\mathbf{u},0}^*$ is continuous from \mathcal{R} into itself. Moreover, for $\langle\mathbf{r},\mathbf{u}\rangle \geq 0$ we have that

$$\begin{aligned} J_{\mathbf{u},0}^*D_{\mathbf{u}}^*\varphi(\mathbf{r}) &= - \int_{\langle\mathbf{r},\mathbf{u}\rangle}^{+\infty} D_{\mathbf{u}}\varphi(\mathbf{r} + (\tau - \langle\mathbf{r},\mathbf{u}\rangle)\mathbf{u}) d\tau \\ &= - [\varphi(\mathbf{r} + (\tau - \langle\mathbf{r},\mathbf{u}\rangle)\mathbf{u})]_{\tau=\langle\mathbf{r},\mathbf{u}\rangle}^{\infty} \\ &= \varphi(\mathbf{r}). \end{aligned}$$

We get the same result for $\langle\mathbf{r},\mathbf{u}\rangle < 0$, which confirms that $J_{\mathbf{u},0}^*$ is a left inverse of $D_{\mathbf{u}}^*$. □

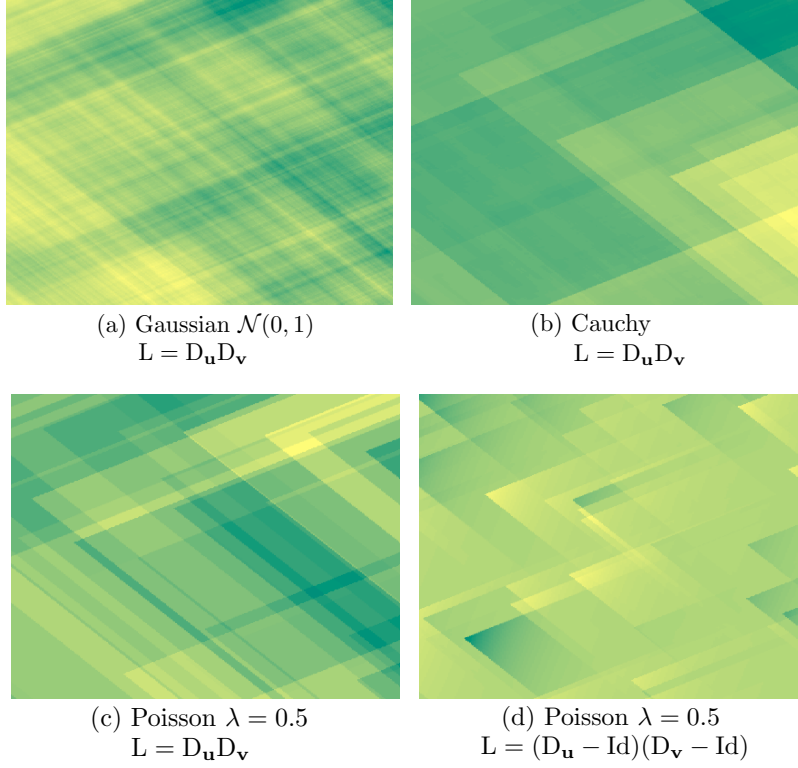


Figure 3: Directional Gaussian or sparse processes

4.2.2 Existence of Directional Sparse Processes

Up to this point, we have covered the first-order directional differential operators. In general, a directional differential operators L can be decomposed as

$$\begin{aligned} L &= (D_{\mathbf{v}_1} - j\omega_1 \text{Id}) \cdots (D_{\mathbf{v}_q} - j\omega_q \text{Id})(D_{\mathbf{u}_1} - \alpha_1 \text{Id}) \cdots (D_{\mathbf{u}_p} - \alpha_p \text{Id}) \\ &= L_{\text{critical}} L_{\text{LSI}}, \end{aligned} \quad (58)$$

where $\Re(\alpha_i) \neq 0$. According to Proposition 7, each of the factors in L_{LSI} has an \mathcal{S} -continuous adjoint left inverse. By composing these operators, we can define a continuous adjoint left inverse L_{LSI}^{*-1} for L_{LSI} . Similarly, the results of Proposition 8 can be employed to form a continuous operator $L_{\text{critical}}^{*-1}$ from \mathcal{S} to \mathcal{R} . Since the constituents of $L_{\text{critical}}^{*-1}$ are not shift-invariant, different composition orders may result in different operators. However, all of them are valid adjoint left inverse operators for L_{critical} .

Finally, $L^{*-1} = L_{\text{critical}}^{*-1} L_{\text{LSI}}^{*-1}$ is a left inverse operator of $L^* = L_{\text{LSI}}^* L_{\text{critical}}^*$, continuous from \mathcal{S} to \mathcal{R} , linear, but not shift-invariant in general. Next, we define generalized stochastic processes based on such L .

Proposition 9. *Let f be a Lévy exponent with $V \in \mathcal{M}(0^+, 2)$, let L be a directional differential operator, and let L^{*-1} stand for its adjoint left inverse. Then, there exists a generalized stochastic process s on \mathcal{S} such that*

$$\widehat{\mathcal{P}}_s(\varphi) = \exp \left(\int_{\mathbb{R}^d} f(L^{*-1} \varphi(\mathbf{r})) d\mathbf{r} \right). \quad (59)$$

The resulting process s is called a directional sparse process (a directional Gaussian process, respectively) if $V \neq 0$ (if $V \equiv 0$, respectively).

Proof. Let $0 < \epsilon \leq 1$ such that $V \in \mathcal{M}(\epsilon, 2)$. As mentioned earlier, L^{*-1} is continuous from \mathcal{S} to \mathcal{R} and, therefore, from \mathcal{S} to $L^\epsilon \cap L^2$. We can now apply Theorem 5 with $p_{\min} = \epsilon$ and $p_{\max} = 2$ to complete the proof. \square

In summary, for all directional differential operators L , we can define the process $s = L^{-1}w$ if V is a Lévy-Schwartz measure. For instance, we can define the classical one-dimensional Lévy processes

(with the point of view of generalized stochastic processes) with $L = D$ as in [UTS11]. We can also define the d -dimensional Mondrian process with $L = D_1 \cdots D_d$ and $V \in \mathcal{M}(0^+, 0)$ which corresponds to a Poisson innovation process (see Table 2), as was done in [UT11] for $d = 2$.

Let $d = 2$. We consider in Figure 3 the case $L = (D_{\mathbf{u}} - \alpha \text{Id})(D_{\mathbf{v}} - \beta \text{Id})$ for some real numbers α, β and vectors $\mathbf{u} = (2, 1)$ and $\mathbf{v} = (2, -1)$. Dark and bright colors indicate large and small values in the simulated realizations, respectively. Note that the first three processes are non-stationary due to the non shift-invariance of the left inverse of $D_{\mathbf{u}}D_{\mathbf{v}}$.

Acknowledgements

The authors are grateful to R. Dalang for the fruitful discussions we had and V. Uhlmann for her helpful suggestions during the writing. The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 267439.

References

- [AJ12] D. Alpay and P.E.T. Jorgensen. Stochastic processes induced by singular operators. *Numerical Functional Analysis and Optimization*, 33(7-9):708–735, 2012.
- [AKBU13] A. Amini, U.S. Kamilov, E. Bostan, and M. Unser. Bayesian estimation for continuous-time sparse stochastic processes. *IEEE Transactions on Signal Processing*, 61(4):907–920, February 15, 2013.
- [AUM11] A. Amini, M. Unser, and F. Marvasti. Compressibility of deterministic and random infinite sequences. *IEEE Transactions on Signal Processing*, 59(11):5193–5201, November 2011.
- [BEK10] H. Biermé, A. Estrade, and I. Kaj. Self-similar random fields and rescaled random balls models. *Journal of Theoretical Probability*, 23(4):1110–1141, 2010.
- [BFKU13] E. Bostan, J. Fageot, U.S. Kamilov, and M. Unser. Map estimators for self-similar sparse stochastic models. In *Proceedings of the Tenth International Workshop on Sampling Theory and Applications (SampTA’13), Bremen, Germany*, 2013.
- [BKNU13] E. Bostan, U.S. Kamilov, M. Nilchian, and M. Unser. Sparse stochastic processes and discretization of linear inverse problems. *IEEE Transactions on Image Processing*, in press, October 2013.
- [BKU12] E. Bostan, U.S. Kamilov, and M. Unser. Reconstruction of biomedical images and sparse stochastic modeling. In *Proceedings of the Ninth IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI’12)*, pages 880–883, Barcelona, Spain, May 2-5, 2012.
- [Bon01] J.-M. Bony. *Cours d’analyse—Théorie des distributions et analyse de Fourier*. Les éditions de l’Ecole Polytechnique, 2001.
- [BORV98] L. Bel, G. Oppenheim, L. Robbiano, and M.C. Viano. Linear distribution processes. *International Journal of Stochastic Analysis*, 11(1):43–58, 1998.
- [CRT06] E.J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.
- [Don06] D.L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- [GCD12] R. Gribonval, V. Cevher, and M.E. Davies. Compressible distributions for high-dimensional statistics. *Information Theory, IEEE Transactions on*, 58(8):5016–5034, 2012.

- [Gel55] I. M. Gelfand. Generalized random processes. *Dokl. Akad. Nauk. SSSR*, 100:853–856, 1955.
- [GV64] I. Gelfand and N. Y. Vilenkin. *Generalized functions. Applications of harmonic analysis*, volume 4. Academic press, New York, USA, 1964.
- [KPAU13] U.S. Kamilov, P. Pad, A. Amini, and M. Unser. MMSE estimation of sparse Lévy processes. *IEEE Transactions on Signal Processing*, 61(1):137–147, January 1, 2013.
- [Mal99] S. Mallat. *A wavelet tour of signal processing*. Academic press, 1999.
- [MG01] D. Mumford and B. Gidas. Stochastic models for generic images. *Quarterly of Applied Mathematics*, 59(1):85–112, 2001.
- [MN68] B.B. Mandelbrot and J.W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422–437, 1968.
- [PU13] P. Pad and M. Unser. On the optimality of operator-like wavelets for sparse ar(1) processes. *Proceedings of the Thirty-Eight IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP’13)*, 59(8):5063–5074, August 2013.
- [Rud91] W. Rudin. *Functional analysis*. International series in pure and applied mathematics, McGraw-Hill, Inc., New York, 1991.
- [Sat94] K.-I. Sato. *Lévy processes and infinitely divisible distributions*. Chapman & Hall, 1994.
- [Sch66] L. Schwartz. *Théorie des distributions*. Hermann, 1966.
- [SLSZ03] A. Srivastava, A.B. Lee, E.P. Simoncelli, and S.C. Zhu. On advances in statistical modeling of natural images. *Journal of Mathematical Imaging and Vision*, 18(1):17–33, 2003.
- [ST94] G. Samorodnitsky and M.S. Taqqu. *Stable non-Gaussian processes: Stochastic models with infinite variance*. Chapman and Hall, 1994.
- [SU12] Q. Sun and M. Unser. Left-inverses of fractional Laplacian and sparse stochastic processes. *Advances in Computational Mathematics*, 36(3):399–441, 2012.
- [Tre67] F. Trèves. *Topological vector spaces, distributions and kernels*. Academic Press, New York, USA, 1967.
- [TVU09] P.D. Tafti, D. Van De Ville, and M. Unser. Invariances, Laplacian-like wavelet bases, and the whitening of fractal processes. *IEEE Transactions on Image Processing*, 18(4):689–702, April 2009.
- [UT11] M. Unser and P.D. Tafti. Stochastic models for sparse and piecewise-smooth signals. *IEEE Transactions on Signal Processing*, 59(3):989–1006, March 2011.
- [UTAK11] M. Unser, P.D. Tafti, A. Amini, and H. Kirshner. A unified formulation of Gaussian vs. sparse stochastic processes—Part 2: Discrete-domain theory. *arXiv:1108.6152v2 [cs.IT]*, 2011.
- [UTS11] M. Unser, P.D. Tafti, and Q. Sun. A unified formulation of Gaussian vs. sparse stochastic processes—Part 1: Continuous-domain theory. *arXiv:1108.6150*, 2011.